

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 27, 2003, 14:24:01 ; Search time 16096.1 Seconds
(without alignments)
8776.084 Million cell updates/sec

Title: US-09-919-408a-1
Perfect score: 3453
Sequence: 1 GCGGCTGGCTACCGCGGC.....AAAAAAAAAAAAAAAAAAAA 3453

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 2889711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl.*

1: gb.ba.*

2: gb.htg.*

3: gb.in.*

4: gb.om.*

5: gb.ov.*

6: gb.pat.*

7: gb.ph.*

8: gb.pl.*

9: gb.pr.*

10: gb.ro.*

11: gb.sts.*

12: gb.sy.*

13: gb.un.*

14: gb.vi.*

15: em.ba.*

16: em.fun.*

17: em.hum.*

18: em.in.*

19: em.mu.*

20: em.om.*

21: em.or.*

22: em.ov.*

23: em.pat.*

24: em.ph.*

25: em.pl.*

26: em.ro.*

27: em.sts.*

28: em.un.*

29: em.vi.*

30: em.htg.hum.*

31: em.htg.inv.*

32: em.htg.other.*

33: em.htg.mus.*

34: em.htg.pln.*

35: em.htg.rod.*

36: em.htg.man.*

37: em.htg.vrt.*

38: em.sy.*

39: em.htgo.hum.*

40: em.htgo.mus.*

41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	3453	100.0	3453	6	AR005211	AR005211 Sequence
2	3453	100.0	3453	6	AR071703	AR071703 Sequence
3	3453	100.0	3453	6	I25169	I25169 Sequence 1
4	3453	100.0	3453	6	I40600	I40600 Sequence 1
5	3426	99.2	3426	6	AX695519	AX695519 Sequence
6	3426	99.2	3426	10	MUSFLK2	M64689 Mouse flk-2
7	3344	96.8	3521	6	I44732	I44732 Sequence 1
8	3344	96.8	3521	10	MMPLT3	X59398 Mouse Plt3
9	2979	86.3	2979	6	AX695520	AX695520 Sequence
10	2227.2	64.5	3501	6	AR005212	AR005212 Sequence
11	2227.2	64.5	3501	6	AR071704	AR071704 Sequence
12	2227.2	64.5	3501	6	I25170	I25170 Sequence 3
13	2227.2	64.5	3501	6	I40601	I40601 Sequence 3
14	2204.8	63.9	3476	6	I40109	I40109 Sequence 1
15	2203	63.8	3475	6	AX695522	AX695522 Sequence
16	2203	63.8	3475	6	I44733	I44733 Sequence 3
17	2203	63.8	3475	9	HSU02687	U02687 Human growt
18	2191.4	63.5	3422	9	HSFLT3RTK	Z26652 H. sapiens F
19	2081.8	60.3	3120	6	AR262794	AR262794 Sequence
20	2081.8	60.3	3120	6	I44518	I44518 Sequence 22
21	2069.8	59.9	2982	6	AX695523	AX695523 Sequence
22	1206	34.9	1894	6	AR149571	AR149571 Sequence
23	1119.6	32.4	2247	6	BD022236	BD022236 Multi-fun
24	474.6	13.7	152564	10	AC127549	AC127549 Mus muscu
25	470.6	13.6	267290	2	AC134441	AC134441 Mus muscu
26	427.4	12.4	429	10	MUS3RTK	L36163 Mus musculu
27	363.2	10.5	480	9	HUM3RTK	L36162 Homo sapien
28	357.4	10.4	3992	6	AX394288	AX394288 Sequence
29	357.4	10.4	3992	6	AX587822	AX587822 Sequence
30	357.4	10.4	3992	9	HSCFMS	X03663 Human mRNA
31	355.8	10.3	3904	9	BC047521	BC047521 Homo sapi
32	349.8	10.1	3545	5	AF153446	AF153446 Danio rer
33	347	10.0	4588	5	OMY417832	AJ417832 Oncorhyn
34	344.2	10.0	389	10	AY094358	AY094358 Rattus no
35	341.8	9.9	4577	14	FCSSMONC	X01643 Feline sarc
36	340.2	9.9	3828	4	CATFMS	J03149 Cat (F. dome
37	336.8	9.8	4222	4	S76596	S76596 c-kit-recep
38	335.6	9.7	3705	10	RRCKITRTK	X62491 R. rattus mr
39	334.4	9.7	266001	2	AC106082	AC106082 Rattus no
40	333.6	9.7	3691	10	BC043054	BC043054 Mus muscu
41	333.6	9.7	3699	10	BC036343	BC036343 Mus muscu
42	332.6	9.6	2176	4	AF263826	AF263826 Bos tauru
43	332.6	9.6	2176	4	AF263827	AF263827 Bos tauru
44	332.6	9.6	3069	4	BOVCKR	D16680 Bovine mRNA
45	331.4	9.6	2919	6	A82459	A82459 Sequence 47

ALIGNMENTS

RESULT 1	AR005211	AR005211	3453 bp	DNA	linear	PAT 04-DEC-1998
LOCUS	AR005211	Sequence 1	from patent US 5747651.			
DEFINITION	AR005211					
ACCESSION	AR005211					
VERSION	AR005211.1	GI:3966090				
KEYWORDS	Unknown.					
SOURCE	Unknown.					
ORGANISM	Unknown.					
REFERENCE	1 (bases 1 to 3453)					
AUTHORS	Lemischka, I.R.					
TITLE	Antibodies against tyrosine kinase receptor flk-1					
JOURNAL	Patent: US 5747651-A 1 05-MAY-1998;					
FEATURES	Location/Qualifiers					

Db 361 AGCTCCCTGGGCTGCCAGCCGCACTTTGATTTTACAAAAACAGAGGAATCGTTTCCATGGCC 420
Qy 421 ATCTTGAACTGACAGACCCAGGAGGAGATACCTACTCCATATTCAGCGGACGC 480
Db 421 ATCTTGAACTGACAGACCCAGGAGGAGATACCTACTCCATATTCAGCGGACGC 480
Qy 481 GCCAACTACACAGTACTGTTTACAGTGAATGTAAAGATACACAGCTGTATGTCTAAGG 540
Db 481 GCCAACTACACAGTACTGTTTACAGTGAATGTAAAGATACACAGCTGTATGTCTAAGG 540
Qy 541 AGACCTTACTTTAGGAAGATGGAATAACAGGATGCACTGCTGCAATCTCCGAGGGTGT 600
Db 541 AGACCTTACTTTAGGAAGATGGAATAACAGGATGCACTGCTGCAATCTCCGAGGGTGT 600
Qy 601 CCGAGCCCACTGTGAGTGGGTCTCTGAGCTCCACAGGATGCACTGCTGCAATCTCCGAGGGTGT 660
Db 601 CCGAGCCCACTGTGAGTGGGTCTCTGAGCTCCACAGGATGCACTGCTGCAATCTCCGAGGGTGT 660
Qy 661 GGCCTGTCTGTTGTACAGAAAGGAGGAAAGGTAATCTCATGAGTTGTTCCGAAACAGACATC 720
Db 661 GGCCTGTCTGTTGTACAGAAAGGAGGAAAGGTAATCTCATGAGTTGTTCCGAAACAGACATC 720
Qy 721 AGATGCTGTCTAGAAATGCACTGGGCGCGAATGCCAAGCTGTTCCACATAGATCTA 780
Db 721 AGATGCTGTCTAGAAATGCACTGGGCGCGAATGCCAAGCTGTTCCACATAGATCTA 780
Qy 781 RACCGGCTCCTCAGAGCACACTGCCCGAGTTATTCCTGAAAGTGGGGNACCCCTGTGG 840
Db 781 RACCGGCTCCTCAGAGCACACTGCCCGAGTTATTCCTGAAAGTGGGGNACCCCTGTGG 840
Qy 841 ATCAGGTGAAGGCCATCATGTGAACCATGGAATCGGGCTCACCTGGGAGCTGGAAGAC 900
Db 841 ATCAGGTGAAGGCCATCATGTGAACCATGGAATCGGGCTCACCTGGGAGCTGGAAGAC 900
Qy 901 AAGCCCTGGAGAGGCGAGCTACTTTGAGATGAGTACCTTCCACAAACAGGACCATG 960
Db 901 AAGCCCTGGAGAGGCGAGCTACTTTGAGATGAGTACCTTCCACAAACAGGACCATG 960
Qy 961 ATTCGGATCTCTCGCTTCTGCTTCTCGTGGGAGGAAACGACACCGGATATACCC 1020
Db 961 ATTCGGATCTCTCGCTTCTGCTTCTCGTGGGAGGAAACGACACCGGATATACCC 1020
Qy 1021 TGCTCTTCTCAAGCACCCAGCCAGTCAAGCTGCTGCTGACCATCTTAGAAAAGGTTT 1080
Db 1021 TGCTCTTCTCAAGCACCCAGCCAGTCAAGCTGCTGCTGACCATCTTAGAAAAGGTTT 1080
Qy 1081 ATAAACGCTACAGCTCGCAAGAGATGAAATTTGACCCCTACGAAAGTTCTGCTTC 1140
Db 1081 ATAAACGCTACAGCTCGCAAGAGATGAAATTTGACCCCTACGAAAGTTCTGCTTC 1140
Qy 1141 TCAGTCAGGTTTAAAGCGTACCCAGATCCGATGCACTGATCTTCTCAAGCCTCA 1200
Db 1141 TCAGTCAGGTTTAAAGCGTACCCAGATCCGATGCACTGATCTTCTCAAGCCTCA 1200
Qy 1201 TTTCTCTGTAACACAGAGGCTGAGGATGGGTACAGATATCTAAATTTTGGCATCAT 1260
Db 1201 TTTCTCTGTAACACAGAGGCTGAGGATGGGTACAGATATCTAAATTTTGGCATCAT 1260
Qy 1261 AAGAAACAGCCAGGAGGATACATATTTATGAGAAATGATGACGCCAGTTTCAACAAA 1320
Db 1261 AAGAAACAGCCAGGAGGATACATATTTATGAGAAATGATGACGCCAGTTTCAACAAA 1320
Qy 1321 ATGTTTCAAGCCAGGAGGATACATATTTATGAGAAATGATGACGCCAGTTTCAACAAA 1380
Db 1321 ATGTTTCAAGCCAGGAGGATACATATTTATGAGAAATGATGACGCCAGTTTCAACAAA 1380
Qy 1381 GCGTCTCTTCTCTGATGGCTTACCCGCTACCTCTGAGCTGGAAGTGTTCGGAC 1440
Db 1381 GCGTCTCTTCTCTGATGGCTTACCCGCTACCTCTGAGCTGGAAGTGTTCGGAC 1440
Qy 1441 AAATCTCCAAATTCACAGGAGGAAATCCAGAGGAGTTTGGAAATAAAAGGCTAACAGA 1500
Db 1441 AAATCTCCAAATTCACAGGAGGAAATCCAGAGGAGTTTGGAAATAAAAGGCTAACAGA 1500

Qy 1501 AAAGTGTTTGGCCAGTGGGTGTGAGCAGTACTCTCTAAATATGAGTGAGGCGGGAAGGG 1560
Db 1501 AAAGTGTTTGGCCAGTGGGTGTGAGCAGTACTCTCTAAATATGAGTGAGGCGGGAAGGG 1560
Qy 1561 CTTCTGTCAAAATGCTGTGCTCAATTTCTATGGGACAGCTTTCGGAACCATCTTTTAA 1620
Db 1561 CTTCTGTCAAAATGCTGTGCTCAATTTCTATGGGACAGCTTTCGGAACCATCTTTTAA 1620
Qy 1621 AACTCACAGGCGCCCTTCCCTTTTCATCCAAAGACAAATCTCTCTATGACCAATTTGGG 1680
Db 1621 AACTCACAGGCGCCCTTCCCTTTTCATCCAAAGACAAATCTCTCTATGACCAATTTGGG 1680
Qy 1681 CTCCTGTCTCCCTTCAATTTGTTGTTCTCATTTGTTGATCTGCCCAAAATACAAAGCAA 1740
Db 1681 CTCCTGTCTCCCTTCAATTTGTTGTTCTCATTTGTTGATCTGCCCAAAATACAAAGCAA 1740
Qy 1741 TTTAGGTACGAGAGTCAAGTGCAGATGATCCAGGTGATCTGCGGCCCTTGGATACGAGTAC 1800
Db 1741 TTTAGGTACGAGAGTCAAGTGCAGATGATCCAGGTGATCTGCGGCCCTTGGATACGAGTAC 1800
Qy 1801 TTCTACGTTGACTTTCAGGGACTATGAATATGACCTTAAAGTGGGAGTTCCCGAGAGAGAA 1860
Db 1801 TTCTACGTTGACTTTCAGGGACTATGAATATGACCTTAAAGTGGGAGTTCCCGAGAGAGAA 1860
Qy 1861 TTAGGTTTGGGAGGTCTGCGGCTCTGCGGCTTTCGCGGAGGTGATGAACGCCACGCC 1920
Db 1861 TTAGGTTTGGGAGGTCTGCGGCTCTGCGGCTTTCGCGGAGGTGATGAACGCCACGCC 1920
Qy 1921 TATGGCATTTAGTAAACCGGAGTCTCAATTCAGGTGCGGCTGAGTAAAGAGAGAA 1980
Db 1921 TATGGCATTTAGTAAACCGGAGTCTCAATTCAGGTGCGGCTGAGTAAAGAGAGAA 1980
Qy 1981 GCTGACAGCTGTGAAAAGAAAGCTCTCATGTGCGAGCTCAAAATGATGACCCACCTGGGA 2040
Db 1981 GCTGACAGCTGTGAAAAGAAAGCTCTCATGTGCGAGCTCAAAATGATGACCCACCTGGGA 2040
Qy 2041 CACCATGACAAATCGTGAATCTGCTGGGGGATGACACTGTFCAGGGCCAGTGTACTTG 2100
Db 2041 CACCATGACAAATCGTGAATCTGCTGGGGGATGACACTGTFCAGGGCCAGTGTACTTG 2100
Qy 2101 ATTTTGTGAATATTTGTTGCTATGCTGAGCTCTCTCACTACTTAAGAGTAAAGAGAGAA 2160
Db 2101 ATTTTGTGAATATTTGTTGCTATGCTGAGCTCTCTCACTACTTAAGAGTAAAGAGAGAA 2160
Qy 2161 TTTTCACAGGACATGGACAGAGATTTTAAAGAAACATAATTTCAAGTCTTACCTTCTTC 2220
Db 2161 TTTTCACAGGACATGGACAGAGATTTTAAAGAAACATAATTTCAAGTCTTACCTTCTTC 2220
Qy 2221 CAGGCACATTTCAAATTTCCAGCATGCTGCTGCTTCCAGGAGAGTTTCACTTACACCCGCTTG 2280
Db 2221 CAGGCACATTTCAAATTTCCAGCATGCTGCTGCTTCCAGGAGAGTTTCACTTACACCCGCTTG 2280
Qy 2281 GATCAGCTCTCAGGCTTCAATGGGAATTTCAATTTCAATTCGAAAGTGAAGTGAATGAA 2340
Db 2281 GATCAGCTCTCAGGCTTCAATGGGAATTTCAATTTCAATTCGAAAGTGAAGTGAATGAA 2340
Qy 2341 AACCAAGAGAGGCTGGCAGAGAGAGAGAGAGATTTTGAACGTGCTGAGCTTGAAGAC 2400
Db 2341 AACCAAGAGAGGCTGGCAGAGAGAGAGAGAGATTTTGAACGTGCTGAGCTTGAAGAC 2400
Qy 2401 CTCCTTTGCTTTGCTTACCAAGTGGCCCAAGGCTGGAATTTCTTGGAGTTTCAAGTCTGT 2460
Db 2401 CTCCTTTGCTTTGCTTACCAAGTGGCCCAAGGCTGGAATTTCTTGGAGTTTCAAGTCTGT 2460
Qy 2461 GTCCACAGAGACCTGGCAGCAGGAGTGTGTTGTTCAACCCAGGAGAGTGTGTAAGATC 2520
Db 2461 GTCCACAGAGACCTGGCAGCAGGAGTGTGTTGTTCAACCCAGGAGAGTGTGTAAGATC 2520
Qy 2521 TGTGACTTTGAGCTGGCCCGGAGACATCTGAGCGACTCAGCTAGCTCTGCTGAGGGCAAC 2580
Db 2521 TGTGACTTTGAGCTGGCCCGGAGACATCTGAGCGACTCAGCTAGCTCTGCTGAGGGCAAC 2580

Tyrosine kinase receptor human flk-2-specific antibodies
Patent: US 5548065-A 1 20-AUG-1996;

JOURNAL
FEATURES

source
BASE COUNT 947 a 822 c 875 g 809 t
ORIGIN

Query Match 100.0%; Score 3453; DB 6; Length 3453;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3453; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 GCGGCTCTGGCTACCGCGCTCCGGAGGCGCATGCGGGCGTTGGCGCAGCGGAGCGGCGG 60
1 GCGGCTCTGGCTACCGCGCTCCGGAGGCGCATGCGGGCGTTGGCGCAGCGGAGCGGCGG 60
61 CGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
61 CGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
121 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
121 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
181 CCATCATCTGATCCGGAATGGTGGAGGATCCCGAGAGACCTCCAGTGTATCCCGAGGCGC 240
181 CCATCATCTGATCCGGAATGGTGGAGGATCCCGAGAGACCTCCAGTGTATCCCGAGGCGC 240
241 CAGAGTGAAGGAGCGGTATATGAAGCGGCGCATGCGGGCGTGGCGGCTGCTGCTGCTGCT 300
241 CAGAGTGAAGGAGCGGTATATGAAGCGGCGCATGCGGGCGTGGCGGCTGCTGCTGCTGCT 300
301 ACCCTGCAAGTGCAGTCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 360
301 ACCCTGCAAGTGCAGTCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 360
361 AGCTCCCTGGGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 420
361 AGCTCCCTGGGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 420
421 ATCTTGAACGCTGACAGAGACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 480
421 ATCTTGAACGCTGACAGAGACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 480
481 GCGAATACACAGTACTGTTTCAAGTGAATGTAAGAGATACACAGCTGTATGCTCAAGG 540
481 GCGAATACACAGTACTGTTTCAAGTGAATGTAAGAGATACACAGCTGTATGCTCAAGG 540
541 AGACCTTACTTTAGGAAGATGGAAGACCGGAGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
541 AGACCTTACTTTAGGAAGATGGAAGACCGGAGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
601 CCGGAGCGGCGGCTGCGGAGTGGGCTGCTGCGAGTCCCAAGGAAAGCTGTAAGAAGAA 660
601 CCGGAGCGGCGGCTGCGGAGTGGGCTGCTGCGAGTCCCAAGGAAAGCTGTAAGAAGAA 660
661 GCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
661 GCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
721 AGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
721 AGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
781 AACGAGGCTCTCAGAGCAGACCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 840
781 AACGAGGCTCTCAGAGCAGACCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 840
841 ATCAGGCTGAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 900
841 ATCAGGCTGAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 900
901 AAGAGCCCTGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 960

2581 GCACGCTGCGGCTGAAGTGGATGGCACCGAGAGCTTATTGAAGGATCTACACATC 2640
2581 GCACGCTGCGGCTGAAGTGGATGGCACCGAGAGCTTATTGAAGGATCTACACATC 2640
2641 AAGAGTGAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2700
2641 AAGAGTGAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2700
2701 CCTTACCTGGCATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2760
2701 CCTTACCTGGCATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2760
2761 ATGAGCAGGAGCATCTATGCGCACAGAGGAGATATCTTTGTAATGCAATCTGCTGGGCT 2820
2761 ATGAGCAGGAGCATCTATGCGCACAGAGGAGATATCTTTGTAATGCAATCTGCTGGGCT 2820
2821 TTTGACTCAAGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2880
2821 TTTGACTCAAGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2880
2881 GCAGAGGAGAGAGAGCATGTATCAGACATCCATCCATCCATCCATCCATCCATCCATCCATCCAT 2940
2881 GCAGAGGAGAGAGAGCATGTATCAGACATCCATCCATCCATCCATCCATCCATCCATCCATCCAT 2940
2941 CAGCAGAGGCGGCGGCTCAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3000
2941 CAGCAGAGGCGGCGGCTCAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3000
3001 AGAAGTTAGCGAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3060
3001 AGAAGTTAGCGAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3060
3061 AGATTAGCTGCGGCTGAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3120
3061 AGATTAGCTGCGGCTGAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3120
3121 AGATGCTGCTGCGGCTGAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3180
3121 AGATGCTGCTGCGGCTGAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3180
3181 AGATGCTGCTGCGGCTGAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3240
3181 AGATGCTGCTGCGGCTGAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3240
3241 AGCTTGAAGGAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3300
3241 AGCTTGAAGGAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3300
3301 ACCAAACCGGTTTTTCTGCTAAGGAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3360
3301 ACCAAACCGGTTTTTCTGCTAAGGAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3360
3361 AATACTATGTAATCTTCTGCTAAGGAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3420
3361 AATACTATGTAATCTTCTGCTAAGGAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3420
3421 TACTGTAAAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3453
3421 TACTGTAAAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3453

RESULT 3
125169 LOCUS 3453 bp DNA linear PAT 07-OCT-1996
DEFINITION Sequence 1 from patent US 5548065.
ACCESSION 125169
VERSION 125169.1 GI:1605039
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 3453)
AUTHORS Lemischka,I.R.

Db	901	AAAGCCCTGGAGGCGAGCTACTTTGAGATGAGTACCTACTCCAAACAGGACCANTG	960
Qy	961	ATTGCGAATCTCTTGGCCCTTTGTCTTCGTTGGGAAGAAACGACACGGATATTATACACC	1020
Db	961	ATTGCGAATCTCTTGGCCCTTTGTCTTCGTTGGGAAGAAACGACACGGATATTATACACC	1020
Qy	1021	TGCTCTTCTCAAGACACCCGACGAGTCAGCGTTGGTGGCCATCTCTAGAAAAAGGGTTT	1080
Db	1021	TGCTCTTCTCAAGACACCCGACGAGTCAGCGTTGGTGGCCATCTCTAGAAAAAGGGTTT	1080
Qy	1081	ATAAAGCGTACCAGCTCGCAAGAAGATGAAATTGACCCGTACGAAAGCTCTGCTTC	1140
Db	1081	ATAAAGCGTACCAGCTCGCAAGAAGATGAAATTGACCCGTACGAAAGCTCTGCTTC	1140
Qy	1141	TCAGTCAGGTTTAAAGCGTACCCAGAAATCCGATGCACGTCGATCTTCTCTCAAGCCTCA	1200
Db	1141	TCAGTCAGGTTTAAAGCGTACCCAGAAATCCGATGCACGTCGATCTTCTCTCAAGCCTCA	1200
Qy	1201	TTTTCTCTGTGAACAGAGGCGCTGGAGATGGGTACAGCATATCTAAATTTTGGGATCAT	1260
Db	1201	TTTTCTCTGTGAACAGAGGCGCTGGAGATGGGTACAGCATATCTAAATTTTGGGATCAT	1260
Qy	1261	AAGAAACAAGCCAGGAGAGTACATATTTCTATGCAGAAATGATGACGCCAGTTCACCAA	1320
Db	1261	AAGAAACAAGCCAGGAGAGTACATATTTCTATGCAGAAATGATGACGCCAGTTCACCAA	1320
Qy	1321	ATGTTTCA CGCTGAATATTAAGAAAGAAACCTCAAGTGTCTAGCAAAATGCTCAGCCAGCCAG	1380
Db	1321	ATGTTTCA CGCTGAATATTAAGAAAGAAACCTCAAGTGTCTAGCAAAATGCTCAGCCAGCCAG	1380
Qy	1381	CGCTCCTGTTCTCTGATGGCTACCCGCTACCTCTTGGACCTGGAAAGAGTGTCCGGAC	1440
Db	1381	CGCTCCTGTTCTCTGATGGCTACCCGCTACCTCTTGGACCTGGAAAGAGTGTCCGGAC	1440
Qy	1441	AAATCTCCCAATGTCACGAGGAGAAATCCCAAGAGGTTTGGTAATAAAAGGCTACAGA	1500
Db	1441	AAATCTCCCAATGTCACGAGGAGAAATCCCAAGAGGTTTGGTAATAAAAGGCTACAGA	1500
Qy	1501	AAAGTGTTTGGCCAGTGGGTGTGAGCAGTACTCTAAATATGATGAGGCCGCGGAAAGGG	1560
Db	1501	AAAGTGTTTGGCCAGTGGGTGTGAGCAGTACTCTAAATATGATGAGGCCGCGGAAAGGG	1560
Qy	1561	CTTCTGCTCAAAATGCTGTGGTCAAAATCTATGGGACAGTCTTGGAAACCAATCTTTTAA	1620
Db	1561	CTTCTGCTCAAAATGCTGTGGTCAAAATCTATGGGACAGTCTTGGAAACCAATCTTTTAA	1620
Qy	1621	AACTCACAGGCCCTTCCCTTTTCATCCAAAGACAAATCTCTTATGTCGACCAAAATACAAAAGCAA	1680
Db	1621	AACTCACAGGCCCTTCCCTTTTCATCCAAAGACAAATCTCTTATGTCGACCAAAATACAAAAGCAA	1680
Qy	1681	CTCTGCTCTCCCTTCATTTGTTTCTCATTTGATCTGTCGACCAAAATACAAAAGCAA	1740
Db	1681	CTCTGCTCTCCCTTCATTTGTTTCTCATTTGATCTGTCGACCAAAATACAAAAGCAA	1740
Qy	1741	TTTAGGTACAGAGTACGCTGAGATGATCCAGGTGATGGGCCCTCTGATAAACGAGTAC	1800
Db	1741	TTTAGGTACAGAGTACGCTGAGATGATCCAGGTGATGGGCCCTCTGATAAACGAGTAC	1800
Qy	1801	TTCTACGTTGACCTTACGGGACTATGATATGACCTTAAAGTGGGAGTTCCTCGAGAGAGAAC	1860
Db	1801	TTCTACGTTGACCTTACGGGACTATGATATGACCTTAAAGTGGGAGTTCCTCGAGAGAGAAC	1860
Qy	1861	TTTAGGTTTGGGAAGGTCCTGGGGTCTGGCGCTTTTCGGGAGGGTGTATGAAACCCACGGCC	1920
Db	1861	TTTAGGTTTGGGAAGGTCCTGGGGTCTGGCGCTTTTCGGGAGGGTGTATGAAACCCACGGCC	1920
Qy	1921	TATGGCATTAGTAAACCGGAGTCTCAATCTAGGTGGCGGTGAAGATGCTAAAGAGAGAA	1980
Db	1921	TATGGCATTAGTAAACCGGAGTCTCAATCTAGGTGGCGGTGAAGATGCTAAAGAGAGAA	1980
Qy	1981	GCTGACAGCTGTGAAAAAGAGCTCTCATGTGCGGAGCTCAAAATGATGACCCACTGGGA	2040

Db	1981	GCTGACAGCTGTGTAAAAAGAAAGCTCTCTCATGTGGAGCTCAAAATGATGACCCACCTGGGA	204
Qy	2041	CACCATGACAACTCGTGTGAATCTGCTGGGGCATGCACACTGTCTAGGGCCAGTGACTTTG	2100
Db	2041	CACCATGACAACTCGTGTGAATCTGCTGGGGCATGCACACTGTCTAGGGCCAGTGACTTTG	2100
Qy	2101	ATTTTGAATATTGTGTCTATGTGTGACCTCTCAACTAAGAAAGTAAAGAGAGAAG	2160
Db	2101	ATTTTGAATATTGTGTCTATGTGTGACCTCTCAACTAAGAAAGTAAAGAGAGAAG	2160
Qy	2161	TTTTCACAGGCATGGACAGAGATTTTAAAGAAATATTTTCAGTTCCTTACCTACTTTTC	2220
Db	2161	TTTTCACAGGCATGGACAGAGATTTTAAAGAAATATTTTCAGTTCCTTACCTACTTTTC	2220
Qy	2221	CAGGCACATTTCAAATTCAGCATGCTGTTCACGAGAAGTTTCAGTTTACCCGGCCCTTG	2280
Db	2221	CAGGCACATTTCAAATTCAGCATGCTGTTCACGAGAAGTTTCAGTTTACCCGGCCCTTG	2280
Qy	2281	GATCAGCTCTCAGGGTTTCAATGGGAAATTCATTTCTGAAGATGAGATTGAATGAA	2340
Db	2281	GATCAGCTCTCAGGGTTTCAATGGGAAATTCATTTCTGAAGATGAGATTGAATGAA	2340
Qy	2341	AACCAGAGAGGCTGGCAGAGAAGAGAGAGAGATTTGAACGTCTGACGTTTGAAGAC	2400
Db	2341	AACCAGAGAGGCTGGCAGAGAAGAGAGAGAGATTTGAACGTCTGACGTTTGAAGAC	2400
Qy	2401	CTCCTTTGCTTTCGCTTACCAAGTGGCCAAAGGCATGGAATTCCTGGAGTTCAAGTCGTGT	2460
Db	2401	CTCCTTTGCTTTCGCTTACCAAGTGGCCAAAGGCATGGAATTCCTGGAGTTCAAGTCGTGT	2460
Qy	2461	GTCCACAGAGACCTGGCAGCAGGAATGTGTGTCAACCACGGGAAGTGGTGAAGATC	2520
Db	2461	GTCCACAGAGACCTGGCAGCAGGAATGTGTGTCAACCACGGGAAGTGGTGAAGATC	2520
Qy	2521	TGTGACTTTGGACTGGCCCGAGACATCTTGAGCGACCTCCAGCTACGTCGTGAGGGCAAC	2580
Db	2521	TGTGACTTTGGACTGGCCCGAGACATCTTGAGCGACCTCCAGCTACGTCGTGAGGGCAAC	2580
Qy	2581	GCACGGCTCCCGTGAAGTGGATGGCACCCGAGAGCTTATTTGGAAGGATCTACACATC	2640
Db	2581	GCACGGCTCCCGTGAAGTGGATGGCACCCGAGAGCTTATTTGGAAGGATCTACACATC	2640
Qy	2641	AAGAGTGACGTCTGGTCTTACGGCATCTTCTCTGGGAGATATTTTCACTGGGTGTGAAC	2700
Db	2641	AAGAGTGACGTCTGGTCTTACGGCATCTTCTCTGGGAGATATTTTCACTGGGTGTGAAC	2700
Qy	2701	CCTTACCTTGGCAATTCCTGTGACGGCTAATCTTCTAATAACTGAATCAGAGTGGATTAAA	2760
Db	2701	CCTTACCTTGGCAATTCCTGTGACGGCTAATCTTCTAATAACTGAATCAGAGTGGATTAAA	2760
Qy	2761	ATGGAGAGCCATTCTATGSCCAGAGAAGGATATATCTTTGTAATGCAATCCTGCTGGGCT	2820
Db	2761	ATGGAGAGCCATTCTATGSCCAGAGAAGGATATATCTTTGTAATGCAATCCTGCTGGGCT	2820
Qy	2821	TTTGAATCAAGGAAGCGCCATCTTCCCAACCTGACTTCAATTTTGAATGCAATCCTGCTGGGCT	2880
Db	2821	TTTGAATCAAGGAAGCGCCATCTTCCCAACCTGACTTCAATTTTGAATGCAATCCTGCTGGGCT	2880
Qy	2881	GCAGGGCAGAGAAGCATGTATCAGACATCCATCCATCTACCAAAACAGGGCGGCCCT	2940
Db	2881	GCAGGGCAGAGAAGCATGTATCAGACATCCATCCATCTACCAAAACAGGGCGGCCCT	2940
Qy	2941	CAGCAGAGAGCGGGCTCAGAGCCCGCATGTCGCCACAGCGCCAGGTGAAGATTCAAGAGAA	3000
Db	2941	CAGCAGAGAGCGGGCTCAGAGCCCGCATGTCGCCACAGCGCCAGGTGAAGATTCAAGAGAA	3000
Qy	3001	AGAAATTAGCAGAGGCGCTTTGGACCCCGCCACCTAGCAGGCTGTAGACCCGAGAGCCA	3060
Db	3001	AGAAATTAGCAGAGGCGCTTTGGACCCCGCCACCTAGCAGGCTGTAGACCCGAGAGCCA	3060
Qy	3061	AGATTAGCCTCGCTCTGAGGAAGGCCCTTACAGCGCGTTGCTTCGCTGGACTTTTCTCT	3120
Db	3061	AGATTAGCCTCGCTCTGAGGAAGGCCCTTACAGCGCGTTGCTTCGCTGGACTTTTCTCT	3120

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Craniata; Muridae; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
AUTHORS
TITLE
JOURNAL
MORRIS, D.W. and ENGELHARD, E.K.
Novel compositions and methods for cancer
Patent: WO 0308583-A 1146 30-JAN-2003;
SAGE Discovery (US)

BASE COUNT	920 a	822 c	875 q	809 t
100	100	100	100	100
200	200	200	200	200
300	300	300	300	300
400	400	400	400	400
500	500	500	500	500
600	600	600	600	600
700	700	700	700	700
800	800	800	800	800
900	900	900	900	900
1000	1000	1000	1000	1000
1100	1100	1100	1100	1100
1200	1200	1200	1200	1200
1300	1300	1300	1300	1300
1400	1400	1400	1400	1400
1500	1500	1500	1500	1500
1600	1600	1600	1600	1600
1700	1700	1700	1700	1700
1800	1800	1800	1800	1800
1900	1900	1900	1900	1900
2000	2000	2000	2000	2000
2100	2100	2100	2100	2100
2200	2200	2200	2200	2200
2300	2300	2300	2300	2300
2400	2400	2400	2400	2400
2500	2500	2500	2500	2500
2600	2600	2600	2600	2600
2700	2700	2700	2700	2700
2800	2800	2800	2800	2800
2900	2900	2900	2900	2900
3000	3000	3000	3000	3000
3100	3100	3100	3100	3100
3200	3200	3200	3200	3200
3300	3300	3300	3300	3300
3400	3400	3400	3400	3400
3500	3500	3500	3500	3500
3600	3600	3600	3600	3600
3700	3700	3700	3700	3700
3800	3800	3800	3800	3800
3900	3900	3900	3900	3900
4000	4000	4000	4000	4000
4100	4100	4100	4100	4100
4200	4200	4200	4200	4200
4300	4300	4300	4300	4300
4400	4400	4400	4400	4400
4500	4500	4500	4500	4500
4600	4600	4600	4600	4600
4700	4700	4700	4700	4700
4800	4800	4800	4800	4800
4900	4900	4900	4900	4900
5000	5000	5000	5000	5000
5100	5100	5100	5100	5100
5200	5200	5200	5200	5200
5300	5300	5300	5300	5300
5400	5400	5400	5400	5400
5500	5500	5500	5500	5500
5600	5600	5600	5600	5600
5700	5700	5700	5700	5700
5800	5800	5800	5800	5800
5900	5900	5900	5900	5900
6000	6000	6000	6000	6000
6100	6100	6100	6100	6100
6200	6200	6200	6200	6200
6300	6300	6300	6300	6300
6400	6400	6400	6400	6400
6500	6500	6500	6500	6500
6600	6600	6600	6600	6600
6700	6700	6700	6700	6700
6800	6800	6800	6800	6800
6900	6900	6900	6900	6900
7000	7000	7000	7000	7000
7100	7100	7100	7100	7100
7200	7200			

Query Match	99.2%	Score 3426;	DB 6;	Length 3426;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 3426;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	CGGCGCTGGCTACCGCGCGCTCCGAGCGCCATGCGGCGTTGGCGCAGCGCAGCGACCGG	60	
Db	1	CGGCGCTGGCTACCGCGCGCTCCGAGCGCCATGCGGCGTTGGCGCAGCGCAGCGACCGG	60	
Qy	61	CGGCTGCTGCTGCTGTTGTTTGTCTCAGTAATGATTTCTTGAGACCGTTTACAAACCAAGAC	120	
Db	61	CGGCTGCTGCTGCTGTTGTTTGTCTCAGTAATGATTTCTTGAGACCGTTTACAAACCAAGAC	120	
Qy	121	CTGCGCTGTGATCAAAGTGTGTTTTTAATCAGTCATGAGAACATGCTCATCAGCGGGAAG	180	
Db	121	CTGCGCTGTGATCAAAGTGTGTTTTTAATCAGTCATGAGAACATGCTCATCAGCGGGAAG	180	
Qy	181	CCATCATGCTACCGAATGGTCGGAGGATCCCCAGAGACCTCCAGTGATACCCCGAGGCGC	240	
Db	181	CCATCATGCTACCGAATGGTCGGAGGATCCCCAGAGACCTCCAGTGATACCCCGAGGCGC	240	
Qy	241	CAGAGTGAAGGACCGGTATATGAAGCGGCCACCGTGGAGGTGGCGAGTCTGGGTCCATC	300	
Db	241	CAGAGTGAAGGACCGGTATATGAAGCGGCCACCGTGGAGGTGGCGAGTCTGGGTCCATC	300	
Qy	301	ACCTCTGAAGTGCAGCTCGCCACCCCGAGGGAGCTTTCTGCTCTGGGTCTTTTAAGCAC	360	
Db	301	ACCTCTGAAGTGCAGCTCGCCACCCCGAGGGAGCTTTCTGCTCTGGGTCTTTTAAGCAC	360	
Qy	361	AGTCCCTGGGCTGCCAGCGCACCTTGATTTTACAAAACAGAGGAATCGTTTCCATGCCC	420	
Db	361	AGTCCCTGGGCTGCCAGCGCACCTTGATTTTACAAAACAGAGGAATCGTTTCCATGCCC	420	
Qy	421	ATCTTTGAACGTGCACAGAGACCCAGGCGAGAGAAATACCTACTCCATATTCAGAGCGAACGC	480	
Db	421	ATCTTTGAACGTGCACAGAGACCCAGGCGAGAGAAATACCTACTCCATATTCAGAGCGAACGC	480	
Qy	481	GCCAACTACAGTACTGTGTTCAAGTGAATGTAAAGATACACAGCTGTATGTGCTTAAGG	540	
Db	481	GCCAACTACAGTACTGTGTTCAAGTGAATGTAAAGATACACAGCTGTATGTGCTTAAGG	540	
Qy	541	AGACCTTACTTTAGGAAGATGGAACACAGGATGCATGCTGTGCATCTCCGAGGGTGT	600	
Db	541	AGACCTTACTTTAGGAAGATGGAACACAGGATGCATGCTGTGCATCTCCGAGGGTGT	600	
Qy	601	CCGGAGCCCACTGTGGAGTGGGTGCTCTGCAGCTCCCAAGGAAAGCTGTAAAGAAAGAA	660	
Db	601	CCGGAGCCCACTGTGGAGTGGGTGCTCTGCAGCTCCCAAGGAAAGCTGTAAAGAAAGAA	660	
Qy	661	GGCCCTGCTGTTGTCAGAAAGGAGGAAAGGTACTTTCATGAGTTGTCGGAACAGACATC	720	
Db	661	GGCCCTGCTGTTGTCAGAAAGGAGGAAAGGTACTTTCATGAGTTGTCGGAACAGACATC	720	
Qy	721	AGATGCTGTGCTAGAAATGCACCTGGGCGCGAATGCACCAAGCTGTTTACCATAGATCTA	780	
Db	721	AGATGCTGTGCTAGAAATGCACCTGGGCGCGAATGCACCAAGCTGTTTACCATAGATCTA	780	
Qy	781	AACGAGGCTCCTCAGAGCACATGCCCCAGTTATTTCTCTGAAAGTGGGGGAACTTGTGG	840	

```
QY 1921 TATGGCAATTAGTAAACCGGAGTCTCAATTCAGTGGCGGTGAAGATGCTAAAGAGAAA 1980
Db 1921 TATGGCAATTAGTAAACCGGAGTCTCAATTCAGTGGCGGTGAAGATGCTAAAGAGAAA 1980
QY 1981 GCTGACAGCTGTGAAAAGAGCTCTCATGTCGGAGCTCAAAATGATGACCCACCTGGGA 2040
Db 1981 GCTGACAGCTGTGAAAAGAGCTCTCATGTCGGAGCTCAAAATGATGACCCACCTGGGA 2040
QY 2041 CACCATGACAAACATCTGTCGATCTGTCGGGCGATGACACTGTGTCAGGCGCAGTGTACTTG 2100
Db 2041 CACCATGACAAACATCTGTCGATCTGTCGGGCGATGACACTGTGTCAGGCGCAGTGTACTTG 2100
QY 2101 ATTTTGAATATTTGTTGCTATGTCGATCTCTCAACTTACCTAAAGAGTAAAGAGAGAAG 2160
Db 2101 ATTTTGAATATTTGTTGCTATGTCGATCTCTCAACTTACCTAAAGAGTAAAGAGAGAAG 2160
QY 2161 TTTCACAGACATGGAACAGAGATTTTAAAGAAATTAAGGAAATTAAGTCTTACCCCTACTTTC 2220
Db 2161 TTTCACAGACATGGAACAGAGATTTTAAAGAAATTAAGGAAATTAAGTCTTACCCCTACTTTC 2220
QY 2221 CAGGCACATTCAAATTCAGCAGTCTGTCGATCTCTCAACTTACCTAAAGAGTAAAGAGAGAAG 2280
Db 2221 CAGGCACATTCAAATTCAGCAGTCTGTCGATCTCTCAACTTACCTAAAGAGTAAAGAGAGAAG 2280
QY 2281 GATCAGCTCTCAGGCTTCAATGGAATTCATTCATCTGAGATGAGATTTGAATATGAA 2340
Db 2281 GATCAGCTCTCAGGCTTCAATGGAATTCATTCATCTGAGATGAGATTTGAATATGAA 2340
QY 2341 AACGAGAGAGCTGTCAGAGAGAGAGAGAGATTTGAACGTCGTCGATTTGAAGAC 2400
Db 2341 AACGAGAGAGCTGTCAGAGAGAGAGAGAGATTTGAACGTCGTCGATTTGAAGAC 2400
QY 2401 CTCCTTTGCTTTGCTTACCAAGTGGCCAAAGGATGGAATTCCTGAGTTCAGTCTGTT 2460
Db 2401 CTCCTTTGCTTTGCTTACCAAGTGGCCAAAGGATGGAATTCCTGAGTTCAGTCTGTT 2460
QY 2461 GTCCACAGAGACTGTCAGCAGGAGTGTGTTGTCACCCAGGAGAGTGTGAGATC 2520
Db 2461 GTCCACAGAGACTGTCAGCAGGAGTGTGTTGTCACCCAGGAGAGTGTGAGATC 2520
QY 2521 TGTGACTTTGGAATCTGCGCCGAGACATCTCTGAGCAGTCCAGTCTGTCAGGCGCAAC 2580
Db 2521 TGTGACTTTGGAATCTGCGCCGAGACATCTCTGAGCAGTCCAGTCTGTCAGGCGCAAC 2580
QY 2581 GCACGCTGCGCGTCAAGTGGATGTCACCCAGGAGTATTTGAAGGAGTACACAATC 2640
Db 2581 GCACGCTGCGCGTCAAGTGGATGTCACCCAGGAGTATTTGAAGGAGTACACAATC 2640
QY 2641 AAGAGTGACGCTGTCGTCACGCACTCTCTCTGAGATATTTTCACTGGGTGTGAAC 2700
Db 2641 AAGAGTGACGCTGTCGTCACGCACTCTCTCTGAGATATTTTCACTGGGTGTGAAC 2700
QY 2701 CTTTACCTGCGATTCCTGTCGACGCTAACTTCTATAAATGATTCAGAGTGGATTTAAA 2760
Db 2701 CTTTACCTGCGATTCCTGTCGACGCTAACTTCTATAAATGATTCAGAGTGGATTTAAA 2760
QY 2761 ATGAGAGAGCATTCATGTCACAGAGAGATATTTGATGATGATTCATCTCTGTCGGCT 2820
Db 2761 ATGAGAGAGCATTCATGTCACAGAGAGATATTTGATGATGATTCATCTCTGTCGGCT 2820
QY 2821 TTTGACTCAAGGAGCGGCTTCTTCCCAACCTGACTTCAATTTTATAGGATGTCAGCTG 2880
Db 2821 TTTGACTCAAGGAGCGGCTTCTTCCCAACCTGACTTCAATTTTATAGGATGTCAGCTG 2880
QY 2881 GCAGGAGAGAGAGCATGATCAGAAATTCATCTTACCAAAACAGCGCGCCCT 2940
Db 2881 GCAGGAGAGAGAGCATGATCAGAAATTCATCTTACCAAAACAGCGCGCCCT 2940
QY 2941 CAGCAGAGAGCGGCTCAGAGCGGCTGCGCAGAGCGGCTGAGGATTCACAGAGAA 3000
Db 2941 CAGCAGAGAGCGGCTCAGAGCGGCTGCGCAGAGCGGCTGAGGATTCACAGAGAA 3000
```

```
QY 3001 AGAAGTTAGCAGGAGGCTTGGACCCCGCCACCTAGCAGGCTGTAGACCGCAGAGCCA 3060
Db 3001 AGAAGTTAGCAGGAGGCTTGGACCCCGCCACCTAGCAGGCTGTAGACCGCAGAGCCA 3060
QY 3061 AGATTAGCCTCGCTCTGAGGAGGCGCCCTACAGCGCGTTCCTCGCTGGACTTTTCTCT 3120
Db 3061 AGATTAGCCTCGCTCTGAGGAGGCGCCCTACAGCGCGTTCCTCGCTGGACTTTTCTCT 3120
QY 3121 AGATGCTGTCGCCATTAATCCTCAAGTGAATCTTATATAAATCAAACTCTCTCGCAGAG 3180
Db 3121 AGATGCTGTCGCCATTAATCCTCAAGTGAATCTTATATAAATCAAACTCTCTCGCAGAG 3180
QY 3181 GCGGAGAGCCCAATTAATGAGCTTGTGTCGAGCGCCCTACCTGGGGCTTTTCACG 3240
Db 3181 GCGGAGAGCCCAATTAATGAGCTTGTGTCGAGCGCCCTACCTGGGGCTTTTCACG 3240
QY 3241 AGCTTGAGGAGGAGCCATGATCTGAATATAGTATATTTGTAATACTACGTGAACAA 3300
Db 3241 AGCTTGAGGAGGAGCCATGATCTGAATATAGTATATTTGTAATACTACGTGAACAA 3300
QY 3301 ACCAAACCGCTTTTGTGTAAGGAAAGCTAAATATGATTTTAAATACTATGTTTAA 3360
Db 3301 ACCAAACCGCTTTTGTGTAAGGAAAGCTAAATATGATTTTAAATACTATGTTTAA 3360
QY 3361 ATATCTATGTAATCTTTTTCATCTATTTAGTATATTTTATGATGGAATAAATCTTC 3420
Db 3361 ATATCTATGTAATCTTTTTCATCTATTTAGTATATTTTATGATGGAATAAATCTTC 3420
QY 3421 TACTCT 3426
Db 3421 TACTCT 3426
```

RESULT 6

```
MUSFLK2
LOCUS Mouse flk-2 mRNA, complete cds. 3426 bp mRNA linear ROD 27-APR-1993
DEFINITION M64689
ACCESSION M64689.1 GI:193327
VERSION 1
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 3426)
AUTHORS Matthews,W., Jordan,C.T., Wiegand,G.W., Pardoll,I.D. and
Lemischka,I.R.
TITLE A receptor tyrosine kinase specific to hematopoietic stem and
progenitor cell-enriched populations
CELL 65 (7), 1143-1152 (1991)
JOURNAL 91292518
MEDLINE PUBMED 1648448
COMMENT Original source text: Mus musculus (strain C3H/He) cDNA to mRNA.
FEATURES
Location/Qualifiers
1..3426
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C3H/He"
/db_xref="taxon:10090"
1..3426
/gene="flk-2"
31..3009
/gene="flk-2"
/codon_start=1
/protein_id="AAA37634.1"
/db_xref="GI:193328"
/translation="MRALQSRDRRLRLVLSVILETVNQLPVKVLISHENN
GSAGKPSRVMRQSPEDLQCTPRQSEGTVEATVEAESGSIQLQVATPGDL
SCLWFKHSLGQCPHFQDQNRGIVSNALNVTVEAGVLLHQSRANTVLTFTVN
VRDTQLVLRPYPFRKMNQDALLCISEGPEPTVEWVLCSSHESCKECPAVRKE
EKVLEHFGTDIRCCARNALGRCTKLPFTIDLNOAPOSTLPOLFKVGEPLWIRKAI
HVNHGFLTWELDKALEGSEYFEMSTYSTRNRMIRILLAFVSVGRNDTGYTCSSS
KHFQSALVLTLEKGFINATSQEYEIDPYEKFCFSVRKAYPRIRCTWFSQSPF
```

source

```
gene
CDS
```


Db	3001	AGAAAGTTAGCAGGAGGCGCTTGGACCCCGCCACCCCTAGCAGGCTGTAGACCGCAGAGCCA	306
Qy	3061	AGATTAGCTCGCCTCTGAGGAAGCGCCTACAGCGCGTTCGCTCGCTGGA	3120
Db	3061	AGATTAGCTCGCCTCTGAGGAAGCGCCTACAGCGCGTTCGCTCGCTGGA	3120
Qy	3121	AGATGCTGTCTGCGCAATTA	3180
Db	3121	AGATGCTGTCTGCGCAATTA	3180
Qy	3181	GC	3240
Db	3181	GC	3240
Qy	3241	AGCTT	3300
Db	3241	AGCTT	3300
Qy	3301	ACCA	3360
Db	3301	ACCA	3360
Qy	3361	ATACTATGTA	3420
Db	3361	ATACTATGTA	3420
Qy	3421	TACTGT	3426
Db	3421	TACTGT	3426
RESULT 7	I44732	I44732	
LOCUS	I44732	Sequence 1 from patent US 5635388.	
DEFINITION	I44732	3521 bp	DNA
ACCESSION	I44732		linear
VERSION	I44732.1	GI:2469445	PAT 07-OCT-1997
KEYWORDS	Unknown.		
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	1 (bases 1 to 3521)		
AUTHORS	Bennett, B.D., Broz, S.D., Matthews, W. and Zeigler, F.C.		
TITLE	Agonist antibodies against the flk2/flt3 receptor and uses thereof		
JOURNAL	Patent: US 5635388-A 1 03-JUN-1997;		
FEATURES	Location/Qualifiers		
source	1..3521		
BASE COUNT	943 a	854 c	907 g
ORIGIN			817 t
Query Match	96.8%	Score 3344;	DB 6; Length 3521;
Best Local Similarity	98.5%	Pred. No. 0;	
Matches 3418, Conservative	0; Mismatches 20; Indels 32; Gaps 3;		
Qy	1	GC	60
Db	52	GC	111
Qy	61	CG	120
Db	112	CG	171
Qy	121	CT	180
Db	172	CT	231
Qy	181	CA	240
Db	232	CC	291
Qy	241	CAG	300

1372	ATGTTACGCTGAATATAAGAAAGAAACCTCAAGTCTAGCAAAATGCTCAGCCAGC	1431
1381	GGGTCTCTGTTCTCTGATCGCTACCGCTACCTCTTGGACCTGGAAGAGTGTTCGGAC	1440
1432	GGGTCTCTGTTCTCTGATGCTTACCGCTACCTCTTGGACCTGGAAGAGTGTTCGGAC	1491
1441	AAATCTCCCAATTCGACCGAGGAAATCCAGAAAGGATTTGGAATAAAGAGCTTACAGA	1500
1492	AAATCTCCCAATTCGACCGAGGAAATCCAGAAAGGATTTGGAATAAAGAGCTTACAGA	1551
1501	AAAGTGTGGCCAGTGGGTGTCGAGCTACTCTAAATATAGTGAAGGCGGGAAGGG	1560
1552	AAAGTGTGGCCAGTGGGTGTCGAGCTACTCTAAATATAGTGAAGGCGGGAAGGG	1611
1561	CTTCTCGTCAAAATGCTGTGCTACAAATTCATGGGCACTCTTTCGCAAAACCATCTTTTA	1620
1612	CTTCTCGTCAAAATGCTGTGCTACAAATTCATGGGCACTCTTTCGCAAAACCATCTTTTA	1671
1621	AACTCAACGAGCCCTTCCCTTTTATCCAAAGAAATCTCTCTTCTATGACCATTTGGG	1680
1672	AACTCAACGAGCCCTTCCCTTTTATCCAAAGAAATCTCTCTTCTATGACCATTTGGG	1731
1681	CTCTGTCTCCCTTCTATGTTGTTCTCATTTGTTGTTGATCTGCCCAAAATACAAAGCAA	1740
1732	CTCTGTCTCCCTTCTATGTTGTTCTCATTTGTTGTTGATCTGCCCAAAATACAAAGCAA	1791
1741	TTTAGTACAGAGTCACTGCTGATGATCCAGTGTGCTGCCCAAAATACAAAGCAA	1800
1792	TTTAGTACAGAGTCACTGCTGATGATCCAGTGTGCTGCCCAAAATACAAAGCAA	1851
1801	TTCTACGTTGAGTTCAGGAGTATGAATATGACCTTAAGTGGGAGTTCCTGAGAGAGAA	1860
1852	TTCTACGTTGAGTTCAGGAGTATGAATATGACCTTAAGTGGGAGTTCCTGAGAGAGAA	1911
1861	TTAGAGTTTGGGAAGTCTCTGGGCTTCTGGGAGGTGATGAACGCCACGCGC	1920
1912	TTAGAGTTTGGGAAGTCTCTGGGCTTCTGGGAGGTGATGAACGCCACGCGC	1971
1921	TATGGCATTAGTAAGAAAGGAGTCTCAATTCAGTGGCGGTGAAGATGCTAAAGAGAA	1980
1972	TATGGCATTAGTAAGAAAGGAGTCTCAATTCAGTGGCGGTGAAGATGCTAAAGAGAA	2031
1981	GCTGACGCTGTGAAAAAGAGTCTCTCATGTGGGAGCTCAAAATGATGACCCCTGGGA	2040
2032	GCTGACGCTGTGAAAAAGAGTCTCTCATGTGGGAGCTCAAAATGATGACCCCTGGGA	2091
2041	CACCATGACACATCGTGAATCTGCTGGGGCATGCACATCTGTCAGGGCCAGTGTACTTG	2100
2092	CACCATGACACATCGTGAATCTGCTGGGGCATGCACATCTGTCAGGGCCAGTGTACTTG	2151
2101	ATTTTGAATATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG	2160
2152	ATTTTGAATATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG	2211
2161	TTTTCAGGACATGACGAGAGATTTTAAAGGAACATAATTCAGTCTTCTTACCTTCTTC	2220
2212	TTTTCAGGACATGACGAGAGATTTTAAAGGAACATAATTCAGTCTTCTTACCTTCTTC	2271
2221	CAGGCACATTCAAATTCAGGATGCTGTTTCAAGAGAGTTCAGTTACCCGCCCTTG	2280
2272	CAGGCACATTCAAATTCAGGATGCTGTTTCAAGAGAGTTCAGTTACCCGCCCTTG	2331
2281	GATCAGCTCTAGGGTTCAATTCAGTCTTCAATTCATCTGAGATGAGTGTGAATAGAA	2340
2332	GATCAGCTCTAGGGTTCAATTCAGTCTTCAATTCATCTGAGATGAGTGTGAATAGAA	2391
2341	AACGAGAGGCTGCGAGAGAGAGAGAGATTTGAACCTGCTGAGCTTGAAGAG	2400
2392	AACGAGAGGCTGCGAGAGAGAGAGATTTGAACCTGCTGAGCTTGAAGAG	2451
2401	CTCCTTTGCTTTCGTTACCAAGTGGCCAAAGGATGGAATTCCTGGAGTTCAAGTCTGTT	2460
2452	CTCCTTTGCTTTCGTTACCAAGTGGCCAAAGGATGGAATTCCTGGAGTTCAAGTCTGTT	2511

Db 172 CTGCTGTGATCAAGTGTGTTTAAATCATGTCATGAGAAATATGGCTCATCAGCGGAAAG 231
Qy 181 CCATCATCGTACCGAATGGTGGAGATCCCCAGAGAGACCTCCAGTGTACCCCGAGGGCC 240
Db 232 CCATCATCGTACCGAATGGTGGAGATCCCCAGAGAGACCTCCAGTGTACCCCGAGGGCC 291
Qy 241 CAGAGTGAAGGACCGGTATATGAAGCGGCACCGTGGAGGTGGCGAGTCTGGTCCATC 300
Db 292 CAGAGTGAAGGACCGGTATATGAAGCGGCACCGTGGAGGTGGCGAGTCTGGTCCATC 351
Qy 301 ACCCTGCAAGTGCAGCTGCCACCCAGGGGACCTTCTGCTCTGGTCTTTAAGCAC 360
Db 352 ACCCTGCAAGTGCAGCTGCCACCCAGGGGACCTTCTGCTCTGGTCTTTAAGCAC 411
Qy 361 AGCTCCCTGGGTGCGCAGCGGCACCTTTGATTTACAAAACAGAGGAATCGTTTCCATGGCC 420
Db 412 AGCTCCCTGGGTGCGCAGCGGCACCTTTGATTTACAAAACAGAGGAATCGTTTCCATGGCC 471
Qy 421 ATCTTTGAACGTGACAGAGACCCAGGAGGAGATACCTACTCCATATTCAGAGCGAACGC 480
Db 472 ATCTTTGAACGTGACAGAGACCCAGGAGGAGATACCTACTCCATATTCAGAGCGAACGC 531
Qy 481 GCCAACTACACAGTACTGTTCAAGTGAATGTAAAGATACACAGCTGTATGTCTAAG 540
Db 532 GCCAACTACACAGTACTGTTCAAGTGAATGTAAAGATACACAGCTGTATGTCTAAG 591
Qy 541 AGACCTTACTTTAGGAAGATGGAACCCAGGATGCACTGCTGCACTCCGAGGGGTGT 600
Db 592 AGACCTTACTTTAGGAAGATGGAACCCAGGATGCACTGCTGCACTCCGAGGGGTGT 651
Qy 601 CCGGAGCCCACTGTGAGTGGGTGCTCTGAGCTCCACAGGAAAGCTGTAAAGAGAA 660
Db 652 CCGGAGCCCACTGTGAGTGGGTGCTCTGAGCTCCACAGGAAAGCTGTAAAGAGAA 711
Qy 661 GGCCTCTGCTGTGTGACAGAGAGGAAAGGTACTTCAATGATGTTTGGAAACAGATC 720
Db 712 GGCCTCTGCTGTGTGACAGAGAGGAAAGGTACTTCAATGATGTTTGGAAACAGATC 771
Qy 721 AGATGCTGTCTAGAAATGCACTGGGCGCGAATGCAACCAAGCTGTTCCACATAGATCTA 780
Db 772 AGATGCTGTCTAGAAATGCACTGGGCGCGAATGCAACCAAGCTGTTCCACATAGATCTA 831
Qy 781 AACCAGGCTCTCAGAGCACACTGCCCGCAGTTATCTCTGAAAGTGGGGAAACCTTGTGG 840
Db 832 AACCAGGCTCTCAGAGCACACTGCCCGCAGTTATCTCTGAAAGTGGGGAAACCTTGTGG 891
Qy 841 ATCAGGTGTAAAGCCATCCATGTGAACCATGGATTCGGGGTCACTGGGAGCTGGAAGAC 900
Db 892 ATCAGGTGTAAAGCCATCCATGTGAACCATGGATTCGGGGTCACTGGGAGCTGGAAGAC 951
Qy 901 AAAGCCCTGAGGAGGCGAGCTATTTGAGTGTAGTACTTCCACAAACAGGACCATG 960
Db 952 AAAGCCCTGAGGAGGCGAGCTATTTGAGTGTAGTACTTCCACAAACAGGACCATG 1011
Qy 961 ATTCGGATTTCTTTGGCTTTGTCTCTCCGTGGGAGGAAACGACACCGGATATTACAC 1020
Db 1012 ATTCGGATTTCTTTGGCTTTGTCTCTCCGTGGGAGGAAACGACACCGGATATTACAC 1071
Qy 1021 TGCTCTTCTCTCAAGACCCCGAGCGAGTCAAGCTTGGTGTGACCATCTAGAAAAGGGTTT 1080
Db 1072 TGCTCTTCTCTCAAGACCCCGAGCGAGTCAAGCTTGGTGTGACCATCTAGAAAAGGGTTT 1131
Qy 1081 ATAAACGCTACAGCTCGCAAGAGAGATGAAATTTGACCCGTAGCAAAAGTTCTGCTTC 1140
Db 1132 ATAAACGCTACAGCTCGCAAGAGAGATGAAATTTGACCCGTAGCAAAAGTTCTGCTTC 1191
Qy 1141 TCAGTCAGGTTTAAAGCGTACCCAGCAATCCGATCGAGTGTGATCTCTCAAGCCTCA 1200
Db 1192 TCAGTCAGGTTTAAAGCGTACCCAGCAATCCGATCGAGTGTGATCTCTCAAGCCTCA 1251
Qy 1201 TTTCTTGTGAAACAGAGAGGCTCGAGGATGGGTACAGCATATCTAAATTTTTCGATCAT 1260

Db 1252 TTTCTTGTGAAACAGAGAGGCTCGAGGATGGGTACAGCATATCTAAATTTTTCGATCAT 1311
Qy 1261 AAGAACAGCCAGGAGAGTACATATTTCTATGACAGAAATGATGACGCCAGTTTCAACAAA 1320
Db 1312 AAGAACAGCCAGGAGAGTACATATTTCTATGACAGAAATGATGACGCCAGTTTCAACAAA 1371
Qy 1321 ATGTTCAAGCTGAATATAGAAAGAAACCTCAAGTGTCTAGCAAAATGCTCAGCAGCCAG 1380
Db 1372 ATGTTCAAGCTGAATATAGAAAGAAACCTCAAGTGTCTAGCAAAATGCTCAGCAGCCAG 1431
Qy 1381 GGGTCTGTCTCTGATGGCTACCGCTACCTCTTGGACCTTGGAGAGTGTTCGGAC 1440
Db 1432 GGGTCTGTCTCTGATGGCTACCGCTACCTCTTGGACCTTGGAGAGTGTTCGGAC 1491
Qy 1441 AAATCTCCCAATTTGACAGGAGAAATCCAGAGGAGTTTGGAAATAAAAGGCTTAAACAGA 1500
Db 1492 AAATCTCCCAATTTGACAGGAGAAATCCAGAGGAGTTTGGAAATAAAAGGCTTAAACAGA 1551
Qy 1501 AAAGTGTTTGGCAGTGGGTGTGAGCAGTACTCTAAATATGAGTGAAGCGCGGAAAGGG 1560
Db 1552 AAAGTGTTTGGCAGTGGGTGTGAGCAGTACTCTAAATATGAGTGAAGCGCGGAAAGGG 1611
Qy 1561 CTTCTGGTCAAAATGCTGTGCTACAAATTTCTATGGCAGCTTTGCGAAACCATCTTTTAA 1620
Db 1612 CTTCTGGTCAAAATGCTGTGCTACAAATTTCTATGGCAGCTTTGCGAAACCATCTTTTAA 1671
Qy 1621 AACTACAGGCGCCCTTCCCTTTCATCCAGACAAACATCTCTTATGCGACCATTTGGG 1680
Db 1672 AACTACAGGCGCCCTTCCCTTTCATCCAGACAAACATCTCTTATGCGACCATTTGGG 1731
Qy 1681 CTCGTCTCCCTTCATTTGTTGTTGATCTGCTGCTGCCACAAATACAAAAGCAA 1740
Db 1732 CTCGTCTCCCTTCATTTGTTGTTGATCTGCTGCTGCCACAAATACAAAAGCAA 1791
Qy 1741 TTTAGGTACAGAGTCAAGTGCAGATGATCCAGTGAATCGGCCCTTGGATTAACGAGTAC 1800
Db 1792 TTTAGGTACAGAGTCAAGTGCAGATGATCCAGTGAATCGGCCCTTGGATTAACGAGTAC 1851
Qy 1801 TTTAGGTACAGAGTCAAGTGCAGATGATCCAGTGAATCGGCCCTTGGATTAACGAGTAC 1860
Db 1852 TTTAGGTACAGAGTCAAGTGCAGATGATCCAGTGAATCGGCCCTTGGATTAACGAGTAC 1911
Qy 1861 TTTAGGTACAGAGTCAAGTGCAGATGATCCAGTGAATCGGCCCTTGGATTAACGAGTAC 1920
Db 1912 TTTAGGTACAGAGTCAAGTGCAGATGATCCAGTGAATCGGCCCTTGGATTAACGAGTAC 1971
Qy 1921 TTTAGGTACAGAGTCAAGTGCAGATGATCCAGTGAATCGGCCCTTGGATTAACGAGTAC 1980
Db 1972 TTTAGGTACAGAGTCAAGTGCAGATGATCCAGTGAATCGGCCCTTGGATTAACGAGTAC 2031
Qy 1981 TTTAGGTACAGAGTCAAGTGCAGATGATCCAGTGAATCGGCCCTTGGATTAACGAGTAC 2040
Db 2032 TTTAGGTACAGAGTCAAGTGCAGATGATCCAGTGAATCGGCCCTTGGATTAACGAGTAC 2091
Qy 2041 TTTAGGTACAGAGTCAAGTGCAGATGATCCAGTGAATCGGCCCTTGGATTAACGAGTAC 2100
Db 2092 TTTAGGTACAGAGTCAAGTGCAGATGATCCAGTGAATCGGCCCTTGGATTAACGAGTAC 2151
Qy 2101 TTTAGGTACAGAGTCAAGTGCAGATGATCCAGTGAATCGGCCCTTGGATTAACGAGTAC 2160
Db 2152 TTTAGGTACAGAGTCAAGTGCAGATGATCCAGTGAATCGGCCCTTGGATTAACGAGTAC 2211
Qy 2161 TTTAGGTACAGAGTCAAGTGCAGATGATCCAGTGAATCGGCCCTTGGATTAACGAGTAC 2220
Db 2212 TTTAGGTACAGAGTCAAGTGCAGATGATCCAGTGAATCGGCCCTTGGATTAACGAGTAC 2271
Qy 2221 TTTAGGTACAGAGTCAAGTGCAGATGATCCAGTGAATCGGCCCTTGGATTAACGAGTAC 2280
Db 2272 TTTAGGTACAGAGTCAAGTGCAGATGATCCAGTGAATCGGCCCTTGGATTAACGAGTAC 2331
Qy 2281 TTTAGGTACAGAGTCAAGTGCAGATGATCCAGTGAATCGGCCCTTGGATTAACGAGTAC 2340
Db 2332 TTTAGGTACAGAGTCAAGTGCAGATGATCCAGTGAATCGGCCCTTGGATTAACGAGTAC 2391

QY 631 AGCTCCACAGGAAAGCTGTAAGAGAGAGCCCTGCTGTTGTCAGAAAGGAGGAAAAG 690
 DB 601 AGCTCCACAGGAAAGCTGTAAGAGAGAGCCCTGCTGTTGTCAGAAAGGAGGAAAAG 660
 QY 691 GTACTTTCATGAGTTGTTGCGAAACAGACATCAGATGCTGTCTAGAAATGCACTGGGCGCG 750
 DB 661 GTACTTTCATGAGTTGTTGCGAAACAGACATCAGATGCTGTCTAGAAATGCACTGGGCGCG 720
 QY 751 GAATGACCAAGCTGTTTACCATAGATCTTAAACAGGCTCTCTCAGAGCACTGCCCCAG 810
 DB 721 GAATGACCAAGCTGTTTACCATAGATCTTAAACAGGCTCTCTCAGAGCACTGCCCCAG 780
 QY 811 TTAATTCCTGAAGTGGGGAAACCTTGTGATCAGGTGTAAAGCCATCCATGTGAACCAT 870
 DB 781 TTAATTCCTGAAGTGGGGAAACCTTGTGATCAGGTGTAAAGCCATCCATGTGAACCAT 840
 QY 871 GGATTCGGGCTCACCTGGGAGCTGGAAGACAAAGCCCTGAGGAGGCGCAGCTACTTTGAG 930
 DB 841 GGATTCGGGCTCACCTGGGAGCTGGAAGACAAAGCCCTGAGGAGGCGCAGCTACTTTGAG 900
 QY 931 ATGAGTACCTTACTCCACAAACAGGACCATGATTGGGATTTCTTTGGGCTTTGTGCTTCC 990
 DB 901 ATGAGTACCTTACTCCACAAACAGGACCATGATTGGGATTTCTTTGGGCTTTGTGCTTCC 960
 QY 991 GTGGGAAGGAACGACACCGGATATTACACCTGCTTCTCCTCAAGCACCCAGCCAGTCA 1050
 DB 961 GTGGGAAGGAACGACACCGGATATTACACCTGCTTCTCCTCAAGCACCCAGCCAGTCA 1020
 QY 1051 GCGTTGGTGACCATCTTAGAAAAGGGTTTATAAAGCTACAGCTCGCAAGAGAGATAT 1110
 DB 1021 GCGTTGGTGACCATCTTAGAAAAGGGTTTATAAAGCTACAGCTCGCAAGAGAGATAT 1080
 QY 1111 GAAATTGACCCGTACGAAAAGTTCTGCTTCTCAGTCAAGTTTAAAGCGTACCCAGAAATC 1170
 DB 1081 GAAATTGACCCGTACGAAAAGTTCTGCTTCTCAGTCAAGTTTAAAGCGTACCCAGAAATC 1140
 QY 1171 CGATGACGTGGATCTTCTCAAGCTCATTTCTGTTGTAACAGAGAGGCGCTGGAGAT 1230
 DB 1141 CGATGACGTGGATCTTCTCAAGCTCATTTCTGTTGTAACAGAGAGGCGCTGGAGAT 1200
 QY 1231 GGGTACAGCATATCTAAATTTGCGATCATAGAACCAAGCCAGGAGAGTACATATTTCTAT 1290
 DB 1201 GGGTACAGCATATCTAAATTTGCGATCATAGAACCAAGCCAGGAGAGTACATATTTCTAT 1260
 QY 1291 GCAGAAATGATGACGCCAGTTTCAACAAATGTTTCAAGCTGAATATAAGAAAGAAACCT 1350
 DB 1261 GCAGAAATGATGACGCCAGTTTCAACAAATGTTTCAAGCTGAATATAAGAAAGAAACCT 1320
 QY 1351 CAAGTGTAGCAATGCTCAGCCAGCGCGCTGCTGTTCTCTGATGGCTACCCGCTA 1410
 DB 1321 CAAGTGTAGCAATGCTCAGCCAGCGCGCTGCTGTTCTCTGATGGCTACCCGCTA 1380
 QY 1411 CCCTCTTGGACCTGGAAGAGTGTTCGGAACAAATCTCCCAATTCGACGAGGAAATCCCA 1470
 DB 1381 CCCTCTTGGACCTGGAAGAGTGTTCGGAACAAATCTCCCAATTCGACGAGGAAATCCCA 1440
 QY 1471 GAAGGAGTTTGGAAATAAGAGCTTAAACAGAAAGTTTGGCCAGTGGGTGTCAGAGCAGT 1530
 DB 1441 GAAGGAGTTTGGAAATAAGAGCTTAAACAGAAAGTTTGGCCAGTGGGTGTCAGAGCAGT 1500
 QY 1531 ACTCTAAATATGAGTGGCGCGGAAAGGGCTTCTGGTCAATGCTGTGCGTACAAATTTCT 1590
 DB 1501 ACTCTAAATATGAGTGGCGCGGAAAGGGCTTCTGGTCAATGCTGTGCGTACAAATTTCT 1560
 QY 1591 ATGGGACGCTTGGGAAACCATCTTTTAAACTCACCAGGCGCTTCCCTTTCAATCCAA 1650
 DB 1561 ATGGGACGCTTGGGAAACCATCTTTTAAACTCACCAGGCGCTTCCCTTTCAATCCAA 1620
 QY 1651 GACAAATCTCTTCTATGCGACCAATGCGGCTCTGCTCTCCCTTCAATGTTGTTCTCAT 1710
 DB 1621 GACAAATCTCTTCTATGCGACCAATGCGGCTCTGCTCTCCCTTCAATGTTGTTCTCAT 1680
 QY 1711 GTGTTGATCTGCCAATAACAAAAGCAATTTAGGTACGAGAGTCAAGTGCAGATGATC 1770

DB 1681 GTGTTGATCTGCCACAAATACAAAAGCAATTTAGGTACGAGAGTCAAGTGCAGATGATC 1740
 QY 1771 CAGGTGACTGGCCCCCTTGGATAACGAGTACTTCTACGTTGACTTTCAGGACTATGAATAT 1830
 DB 1741 CAGGTGACTGGCCCCCTTGGATAACGAGTACTTCTACGTTGACTTTCAGGACTATGAATAT 1800
 QY 1831 GACCTTAAAGTGGGAGTTCCCGAGAGAGAACTTAGAGTTTGGAGAGTCTCTGGGCTCTGGC 1890
 DB 1801 GACCTTAAAGTGGGAGTTCCCGAGAGAGAACTTAGAGTTTGGAGAGTCTCTGGGCTCTGGC 1860
 QY 1891 GCTTTTCGGAGGGTGTATGAACCGCCACTATGGCAATTAGTAAACCGGAGTCTCAAT 1950
 DB 1861 GCTTTTCGGAGGGTGTATGAACCGCCACTATGGCAATTAGTAAACCGGAGTCTCAAT 1920
 QY 1951 CAGGTGGCGGTGAAGATGCTTAAAGAGAAAGCTGACAGCTGTGTGAAAAGAGGCTCTCATG 2010
 DB 1921 CAGGTGGCGGTGAAGATGCTTAAAGAGAAAGCTGACAGCTGTGTGAAAAGAGGCTCTCATG 1980
 QY 2011 TCGGAGCTCAAAATGATGACCCCTGGGACACCATGACACATCTGTGAATCTGCTGGG 2070
 DB 1981 TCGGAGCTCAAAATGATGACCCCTGGGACACCATGACACATCTGTGAATCTGCTGGG 2040
 QY 2071 GCATGACACTGTCTCAGGCGCAGTGTACTTGTATTTTGAATTTTGTCTATGCTGACCTC 2130
 DB 2041 GCATGACACTGTCTCAGGCGCAGTGTACTTGTATTTTGAATTTTGTCTATGCTGACCTC 2100
 QY 2131 CTCAACTTACTAAGAAAGTAAAGAGAGAGTTCACAGGACATGGAAGAGATTTTAAAG 2190
 DB 2101 CTCAACTTACTAAGAAAGTAAAGAGAGAGTTCACAGGACATGGAAGAGATTTTAAAG 2160
 QY 2191 GAACATTAATTTTCACTTCTTACCTTCTTCCAGGACATTCATAATTCAGCATGCTCGT 2250
 DB 2161 GAACATTAATTTTCACTTCTTACCTTCTTCCAGGACATTCATAATTCAGCATGCTCGT 2220
 QY 2251 TCAGGAAAGTTTCACTTACACCCCTTGGATCAGCTCTCAGGCTTCAATGGGAATTTCA 2310
 DB 2221 TCAGGAAAGTTTCACTTACACCCCTTGGATCAGCTCTCAGGCTTCAATGGGAATTTCA 2280
 QY 2311 ATTCATTTCTGAAGATGAGATTGAATATGAATAATGAATAATGAATAATGAATAATGA 2370
 DB 2281 ATTCATTTCTGAAGATGAGATTGAATATGAATAATGAATAATGAATAATGAATAATGA 2340
 QY 2371 GAAGATTTGAACGTGCTGACGTTTGAAGACCTCTTGTGCTTTCGTACCAAGTGGCCAAA 2430
 DB 2341 GAAGATTTGAACGTGCTGACGTTTGAAGACCTCTTGTGCTTTCGTACCAAGTGGCCAAA 2400
 QY 2431 GGCATGGAATTTCTGGAGTTTCAAGTCTGCTCCACAGAGACCTGGGACGACGAGGATG 2490
 DB 2401 GGCATGGAATTTCTGGAGTTTCAAGTCTGCTCCACAGAGACCTGGGACGACGAGGATG 2460
 QY 2491 TTGCTCACCCACGGAAGGTTGGAAGATCTGTGACTTTTGAAGTGGCCCGGACATCTCTG 2550
 DB 2461 TTGCTCACCCACGGAAGGTTGGAAGATCTGTGACTTTTGAAGTGGCCCGGACATCTCTG 2520
 QY 2551 AGCAGCTCCAGCTACGCTGTCAGGGGCAACGCAAGCTGCGGCTGAAGTGGATGCAACCC 2610
 DB 2521 AGCAGCTCCAGCTACGCTGTCAGGGGCAACGCAAGCTGCGGCTGAAGTGGATGCAACCC 2580
 QY 2611 GAGAGCTTATTTGAAGGATCTACACATCAAGAGTACGCTGCTCTGAGGATCTCTT 2670
 DB 2581 GAGAGCTTATTTGAAGGATCTACACATCAAGAGTACGCTGCTCTGAGGATCTCTT 2640
 QY 2671 CTCTGGGAGATATTTTCACTGGGTGTGAACCTTACCTTGGCATTTCTGTGACGCTTAC 2730
 DB 2641 CTCTGGGAGATATTTTCACTGGGTGTGAACCTTACCTTGGCATTTCTGTGACGCTTAC 2700
 QY 2731 TTCTATAAAGTGTATGAGTGGATTTTAAATGAGAGAGCAATTTCTATGCTCAAGAGGG 2790
 DB 2701 TTCTATAAAGTGTATGAGTGGATTTTAAATGAGAGAGCAATTTCTATGCTCAAGAGGG 2760
 QY 2791 ATATATTTTGAATGCAATCTCTGCTGGCTTTTGAATCAAGGAGGGGCTTCTTCCC 2850

Db 2761 ATATACCTTTGTAATGCAATCCTGCTGGGCTTTTGAATCAAGGAAGCGGCCATCCTTCCCC 2820
Qy 2851 AACCTGACCTTCATTTTAGATGTACGTGGCAGAGCAGAGCAAGATGTATCAGAACA 2910
Db 2821 AACCTGACCTTCATTTTAGATGTACGTGGCAGAGCAGAGCAAGATGTATCAGAACA 2880
Qy 2911 TCCATCCATCTACCAAAACAGGCGGCCCTCAGCAGAGAGGCGGCTCAGAGCCAGTCG 2970
Db 2881 TCCATCCATCTACCAAAACAGGCGGCCCTCAGCAGAGAGGCGGCTCAGAGCCAGTCG 2940
Qy 2971 CCACAGCGCCAGGTGAAGATTACAGAGAAAGAAAGTTAG 3009
Db 2941 CCACAGCGCCAGGTGAAGATTACAGAGAAAGAAAGTTAG 2979

RESULT 10
AR005212
LOCUS AR005212 3501 bp DNA linear PAT 04-DEC-1998
DEFINITION Sequence 3 from patent US 5747651.
ACCESSION AR005212
VERSION AR005212.1 GI:3966091
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 3501)
AUTHORS Lemischka, I.R.
TITLE Antibodies against tyrosine kinase receptor flk-1
JOURNAL Patent: US 5747651-A 3 05-MAY-1998;
FEATURES
Location/Qualifiers
1..3501
/organism="unknown"
BASE COUNT 1068 a 709 c 784 g 940 t
ORIGIN

Query Match 64.5%; Score 2227.2; DB 6; Length 3501;
Best Local Similarity 79.6%; Pred No. 0;
Matches 2775; Conservative 0; Mismatches 653; Indels 56; Gaps 10;

Qy 8 GGCTACCGCGCGCTCCGAGGCGCATGCGGGCTTGGCGAGCGCAGCGACCGCGGCTGC 67
Db 35 GGGGACCCCGGCGCTCCGAGGCGCATGCGGGCTTGGC---GCGCGACGGGSCACCGTGC 91
Qy 68 TGCTGCTTGTGTTTGTGTAAGTATCTTGACACCGTTACAAACAGACCTGCCTG 127
Db 92 CGCTGCTGCTGTTTGTGTAAGTATCTTGACACCGTTACAAACAGACCTGCCTG 151
Qy 128 TGATCAAGTGTGTTTAAATCAATGATGAGCAATGGCTCATCAGCGGGGAGCCATCAT 187
Db 152 TGATCAAGTGTGTTTAAATCAATGATGAGCAATGGCTCATCAGCGGGGAGTGCATCAT 211
Qy 188 CGTACCGAATGGTGGAGGATCCCGAGAGACCTCCAGTGTACCCCGAGGCGCCAGAGTG 247
Db 212 CATATCCCATGGTATCAGATCCCGAGAGACCTCGGGTGTGGTGTAGAGCCCGAGAGCT 271
Qy 248 AAGGACGGTATATGAAGCGCCACCGTGGAGTGGCGAGTCTGGGTGCATCACCCTGC 307
Db 272 CAGGACAGTGTACGAAGCTCCCGTGTGGAAGTGTATCTGCTTCCATCAGCTGC 331
Qy 308 AAGTGCAGCTGCCACCCCGGAGACCTTCTGCTGCTGCTGCTGCTTAAAGCAGACGCTCC 367
Db 332 AAGTGTGCTGATGCCCGGAGACATTTCTGCTCTGCTGCTTAAAGCAGACGCTCC 391
Qy 368 TGGGCTGCCAGCGGACCTTTGATTTACAAACAGAGGAATCGTTTCCATGGCCATCTTGA 427
Db 392 TGAATTCGACGCCACATTTTGAATTTACAAACAGAGGAGTGTTCATGCTGATTTGA 451
Qy 428 ACGTACAGACACCCAGCAGAGATACCTATCTCCATTTACAGCGGAAACCGGCCAAT 487
Db 452 AAATGACAGAAACCCAGCTGGAGATACCTACTTTTTTATTCAGAGTGAAGTACCAATT 511
Qy 488 ACACAGTACTGTTTACAGTGAATGTAGAGATACACAGCTGTATGTGTAAGGAGCTT 547

Db 512 ACACATATTGTTTACAGTGAGTATAGAAATACCTCGCTTTTACACATTAAGAAGACCTT 571
Qy 548 ACTTTAGGAAGATGGAACACAGGATGCACTGCTCTGCATCTCCGAGGCTGTTCCGAGC 607
Db 572 ACTTTAGAAATATGAAACACAGGACGCCCTGCTGTCATATCTGAGAGCTTCCAGAGC 631
Qy 608 CCACCTGTGAGTGGTCTCTGAGCTCCACAGGGAAGCTGTAAAGAAAGAGGCGCTG 667
Db 632 CGATCGTGAATGGGTGCTTTCGATTTACAGGGGGAAAGCTGTAAAGAAAGAAAGTCCAG 691
Qy 668 CTGTTGTGAGAAAGAGGAAAGAGTACTTTCATGAGTTGTTCCGAAACAGACATCAGATGCT 727
Db 692 CTGTTGTTAAAGAGGAAAGAAAGTCTTTCATGAATTTATGGGACGGACATAGGTGCT 751
Qy 728 GTGCTAGAAATGCATCGGCGCGCAATGCACCAAGCTGTTCCACATAGATCTAAACCCAGG 787
Db 752 GTGCCAGAAATGAATCTGGCAGGGAATGCACAGGCTGTTCAATAGATCTAAATCAA 811
Qy 788 CTCCTCAGAGCAGCTGCCCGAGTTATTCCTGAAAGTGGGGGAACCTTGTGGATCAGGT 847
Db 812 CTCCTCAGACCAATTCGACCAATTTATTTCTTAAAGTAGGGGAACCTTATGATAGGT 871
Qy 848 GTAAGGCCATCCATGTAACCATGATTCGGGCTCACCTGGGAGCTGGGAAGCAAAAGCCC 907
Db 872 GCATAAGCTGTTTCATGTGAACCATGATTCGGGCTCACCTGGGAATTTAGAAAACAAAGCAC 931
Qy 908 TGGAGGAGGCGAGCTACTTTTGAGATGAGTACCTACTCCACAAACAGGACCATGATTCGGA 967
Db 932 TCGAGGAGGCGCACTACTTTTGAGATGAGTACCTATTCAACAAACAGAACTATGATACGA 991
Qy 968 TTCTCTGGCCCTTGTGCTTCCGTTGGGAAGAAACGACACCGGATATTACACCTGCTCTT 1027
Db 992 TTCTGTTGCTTTTGTATCATCAGTGGCAAGAAACGACACCGGATATTACACCTGTTCT 1051
Qy 1028 CCTCAAAGCACCCAGCCAGTCAGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1087
Db 1052 CTTCAAAGCATCCCGAGTCAATCAGCTTTGGTTTACCATCTGAGGAAAGGATTTATAAATG 1111
Qy 1088 CTACGAGCTCCCAAGAGATGATAAATGACCCCTGACGAAAGTTCTGCTTCTCAGTCA 1147
Db 1112 CTACCAATTCAGTGAAGATTTGAAATGACCAATATGAGAGGTTTGTGTTTCTGTC 1171
Qy 1148 GGTTTAAAGCGTACCCAGAAATCCGATGCACTGGATCTTCTCTCAAGCTCATTTCTCT 1207
Db 1172 GGTTTAAAGCGTACCCAGAAATCAGATGTACGTGGACCTTCTCTCGAATAATCATTTCT 1231
Qy 1208 GTGACAGAGAGGCTGGAGATGGGTACAGCATATCAAATTTTGGCATCATAGAAACA 1267
Db 1232 GTGAGCAAAAGGGTCTTGATAAACGGATACAGCATATCCAAAGTTTGGCAATCATAAGCAC 1291
Qy 1268 AGCCAGAGAGTACATATTTCTATGAGAAATGATGACGCCAGTTCCACCAAAATGTTCA 1327
Db 1292 AGCCAGAGATATATATTCATGAGAAATATGATGCCCCAATTTACCAAAATGTTCA 1351
Qy 1328 CGCTGAATATAAGAAAGAAACCTCAAGTGTAGCAAAATGCTCAGCCAGCCAGGCGTCT 1387
Db 1352 CGCTGAATATAAGAAAGAAACCTCAAGTGTCCGAGAAGCATCGCAAGTCAAGGCGTCT 1411
Qy 1388 GTTCTCTGATGGTACCCGCTACCTTGGACCTGGAAGAGTGTTCGGAACAATCTC 1447
Db 1412 GTTCTCGATGAGTACCCATTTACCATCTTGGACCTGGAAGAGTGTTCAGACAAAGTCTC 1471
Qy 1448 CCAATTCAGAGGAAATCCAGAGAGTGTGGAATATAAGAGCTTAAAGAGTAAAGAAAGT 1507
Db 1472 CCACTGCAAGAGAGATCAGAGAGGATCTGGAATAGAAAGCTTAAAGAGTAAAGAGT 1531
Qy 1508 TTGGCAGTGGTGTCCAGCAGTACTCTAAATATGAGTGGGCGCGGAAAGGCTTCTGG 1567
Db 1532 TTGGACAGTGGTGTCCAGCAGTACTCTAAACATGAGTGAAGCCATAAAGGGTCTCTGG 1591
Qy 1568 TCAATGCTGTGCTACAAATCTATGGCAGCTTTCGGAACCATCTTTTTTAACTCAC 1627
Db 1592 TCAAGTCTGTGCATACAAATTTCCCTTGGCACATCTTGTGAGACGATCTTTTTTAACTCTC 1651

QY	1628	CAGGCCCTTCCCTTTCATCCCAAGACAAATCTCCTTCTATGGGACCATTTGGCTCTGTC	1687
Db	1652	CAGGCCCTTCCCTTTCATCCCAAGACAAATCTCCTTCTATGGGACCATTTGGCTCTGTC	1711
QY	1688	TCCCTTCAATGTTGTTCTCATTTGTTGATCTGCCACAAATACAAATAAAGCAATTTAGGT	1747
Db	1712	TCCCTTCAATGTTGTTTAAACCTCTAATTTGTCAAGATACAAATAAAGCAATTTAGGT	1771
QY	1748	ACGAGAGTCAGCTGCAGATGATCCAGGTGACTGGCCCTGGATAACGAGTACTTCTACG	1807
Db	1772	ATGAAGCCAGCTACAGATGGTACAGGTGACCGCTCTCAGATAATGAGTACTTCTACG	1831
QY	1808	TTGACTTTCAGGGACTATGAATATGAATTAAGTGGAGTTCCCGAGAGAACTTAGAGT	1867
Db	1832	TTGATTTTCAGAGAAATGAATATGATCTCAATGGGAGTTTCCAAGAGAAAAATTTAGAGT	1891
QY	1868	TTGGGAAGTCTCGGGTCTGGGCTTTCCGGAGGTGATGAACGCCACGGCTTATGGCA	1927
Db	1892	TTGGGAAGGTACTAGGATCAGGTGCTTTTGGAAAAAGTGATGAACGCAACAGCTTATGGAA	1951
QY	1928	TTAGTAAAAAGCGGAGTCTCAATTCAGGTGGCGGTGAAGATGCTAAAAAGAAAAAGCTGACA	1987
Db	1952	TTAGCAAAAACAGGAGTCTCAATTCAGGTTCGGTCAAAATGCTGAAGAAAAAGCAGACA	2011
QY	1988	GCTGTGAAAAAGAGTCTCATGTTCGGAGCTCAAAATGATGACCCACTTCGGACACCATG	2047
Db	2012	GCTCTGAAGAGAGGCACTCATGTCAAACTCAAGATGATGACCCAGCTGGGAAGCCACG	2071
QY	2048	ACAACATCGTGAATCTGCTGGGGCATGACACTGTGAGGCCCGCTGACTGATTTTGG	2107
Db	2072	AGAATATGTGAACCTGCTGGGGCGTGCACACTGTGAGGACCAATTTACTTGAATTTTG	2131
QY	2108	AATATTGTTGCTATGTTGACTCTCACTCACTACCTTAAGAAGTAAAGAGAGAGTTTTCACA	2167
Db	2132	AATACTGTTGCTATGTTGATCTTCACTCACTATCTAAGAAGTAAAGAGAAAAATTTTACA	2191
QY	2168	GGACATGGACAGAGATTTTAAAGAAACATAATTTTCACTTCTTACCTTCTTCCAGGCAC	2227
Db	2192	GGACTTGGACAGAGATTTTCAAGGAAACACAATTTTCACTTCTTACCTTCTTCCAAATCAC	2251
QY	2228	ATTCAAAATCCAGATCGCTGTTTCAAGAGAGTTTCACTTACCCCGCTTGTGATCAGC	2287
Db	2252	ATCCAAATTCAGCATCGCTGTTTCAAGAGAGTTTCACTTACCCCGCTTGTGATCAGC	2311
QY	2288	TCTCAGGTTCAATGGAATTCATTTCACTTCAAGATGATGAATGAATGAATGAATGAATGA	2347
Db	2312	TCTCAGGCTTCAATGGAATTCATTTCACTTCAAGATGATGAATGAATGAATGAATGAATGA	2371
QY	2348	AGAGGCTGGCAGAGAGAGAGAGATTTGAACGCTGCTGACGTTTGAAGACCTTCTCTTT	2407
Db	2372	AAAGGCTG-----GAAGAGAGAGAGACTTGAATGTGCTTACATTTGAAGATCTTCTTT	2425
QY	2408	GCTTTGGCTACCAAGTGGCCAAAGCATGGAATTCCTGGAGTTCAAGTCTGTGTGTCACA	2467
Db	2426	GCTTTGCATATCAAGTTGCCAAAGCAATGGAATTTCTGGAATTTAAGTCTGTGTGTCACA	2485
QY	2468	GAGACCTGGCCAGAGATGTTGTTGTCACCCAGGAGTGGTGAAGTCTGTGACT	2527
Db	2486	GAGACCTGGCCAGAGATGTTGTTGTCACCCAGGAGTGGTGAAGTATGTGACT	2545
QY	2528	TTGGACTGGCCCGAGACATCTGAGCGACTCCAGCTAGTCTGTGAGGGGCAACGACCGGC	2587
Db	2546	TTGGATTTGGCTCGAGATATATGAGTGATTTCAACTATGTTGTGAGGGGCAATGCCGTC	2605
QY	2588	TGCCGCTGAAGTGGATGGACCCCGAGAGCTTATTTGAAGGATCTTACAAATCAAGATG	2647
Db	2606	TGCCTGTAAATGATGGCCCGGAGCGCTTGTGAAGGCATCTTACACCATTAAGAGTG	2665
QY	2648	ACGTCTGGTCTACGGCATCTTCTCTGGGAGATATTTTCTGAGGAGTGTGAACTTAC	2707
Db	2666	ATGTCTGGTCATAGGAATATTAAGTGGGAAATCTTCTCAGTGGTGTGATCCTTACC	2725
QY	2708	CTGGCATTCCTGTGACGCTAACTTCTATAAACTGATTCAGAGTGGATTTAAATGGAGC	2767
Db	2726	CTGGCATTCCTGTGATGCTAACTTCTATAAACTGATTCAGAGTGGATTTAAATGGAGC	2785
QY	2768	AGCCATTTCTATGCCACAGAGGATATATTTGTAATGCAATCTCTGCTGGCTTTTACT	2827
Db	2786	AGCCATTTTATGCTACAGAGAAATATATTAATGCAATCTCTGCTGGCTTTTACT	2845
QY	2828	CAAGGAGGCGCATCTTCCCACTGACTGATTTTATAGGATGTCAGCTGCAGAGG	2887
Db	2846	CAAGGAGGCGCATCTTCCCTTAATTTGACTTCTGTTTATAGGATGTCAGCTGCAGATG	2905
QY	2888	CAGAGAAAG-----CATGTATCAGAAACATCCATCTCATCTTACC	2924
Db	2906	CAGAGAAAGATGATCAGATGTGGATGGCGCTGTTTCGGAATGCTCTCACACCTTACC	2965
QY	2925	AAAAAGGGGGCCCTCAGCAGAGAGCGGGCTCAGAGC-CCAGTCCCAACAGGCCAGG	2983
Db	2966	AAAAAGGGGACCTTTCAGCAGAGAGATGATTTGGGGTACTCTCTCCGAGGCTCAGG	3025
QY	2984	T-GAAGATTCACAGAGAAAGTTAGTCAGGAGGAGCTTGGACCCCGC-----ACCCTAG	3038
Db	3026	TCAAGATTCGTAGAGGAACAATTTAGTTTAAAGGACTTCCCTCCCTATCCCTATA	3085
QY	3039	CAGGCTGTAGACCGCAGAGCCAGATTTAGCTCGCTCT--GAGGAAGCGCCCTACAGCG	3096
Db	3086	CAGGCTGTAGATTTACCAAAACAAGATTAATTTATCATCTAAAGAAATCTATTTATCAAC	3145
QY	3097	CGTTGCTTCGCTGGACTTTTCTAGATGCTGTCTGCTTACTC-----CAAAGTGA	3149
Db	3146	TGCTGCTTCACAGACTTTTCTCTAGAGCGCTGCTGCTTACTTCTTGTGTTTCAAGGGA	3205
QY	3150	CTTCTATAAATCAAACTCTCTCTGACAGCGGGAGAGCCAAATATAGACTTGTGTTG	3209
Db	3206	CTTTTGTAAATCAAACTCTCTGTCACAAAGCGAGGAGAGCTGATAATGAACCTTTATG	3265
QY	3210	TGAGCCCGCTACCTCGGGGCTTTTCCAGAGCTTTCAGGAGGAGCCATCTATCTGAAA	3269
Db	3266	GAGCATTTGATCTGCATCCAAAGGCTTCTCAGGCGGCTTGTGATGATTTGTGCTGAAG	3325
QY	3270	TATAGTATATTTCTGTAAATACGTCGAAACAAACCAACCCCGTTTTCGTAAGGGAAGC	3329
Db	3326	TACAGTATATTTCTGTAAATACATAAAACAAA-----AGCATTTTGTGTAAGGAGAGC	3378
QY	3330	TAAATATGATTTTAAAAATCTATGTTTAAATACTATGTAACCTTTTCTATCTATTTAG	3389
Db	3379	TAATATGATTTT---AAGTCTATGTTTAAATAATAATGTAAATTTTTCAGCTATTTAG	3436
QY	3390	TGATATTTTATGATGGAATAAACTTTTCTACTGTAAATAAAAAAAAAAAAAAAAAAAAA	3449
Db	3437	TGATATATTTTATGGTGGGAATAAAATTTTCTACTACAGAAAAAAAAAAAAAAAAAAAA	3496
QY	3450	AAAA 3453	
Db	3497	AAAA 3500	
RESULT 11			
AR071704			
LOCUS	AR071704	3501 bp	DNA
DEFINITION	Sequence 3 from patent US 5912133.	linear	PAT 18-FEB-2000
ACCESSION	AR071704		
VERSION	AR071704.1	GI:7222592	
KEYWORDS	Unknown.		
SOURCE	Unknown.		
ORGANISM	Unclassified.		
REFERENCE	1 (bases 1 to 3501)		
AUTHORS	Lemischka, I.R.		
TITLE	Method for isolating stem cells expressing flk-1 receptors		
JOURNAL	Patent: US 5912133-A 3 15-JUN-1999;		
FEATURES	Location/Qualifiers		
	1..3501		

BASE COUNT ORIGIN	1068 a	709 c	784 g	940 t	/organism="unknown"
Query Match	64.5%; Score 2227.2; DB 6; Length 3501;				
Best Local Similarity	79.6%; Pred. No. 0;				
Matches 2775; Conservative	0; Mismatches 653; Indels 56; Gaps 10;				
Qy	8	GGCTACCGCGCTCCGGAGGCGCATGCGGGGCTTGGCGCAGCGAGCGACCGCGGCTGC	67		
Db	35	GGGAGCCCGGCTCCGGAGGCGCATGCGGGGCTTGGCGCAGCGAGCGACCGGTGC	91		
Qy	68	TGCTGCTTGTGTTTGTGCTAGTAATCTTTGAGACCGTTTACAAACCAAGACCTGCGCTG	127		
Db	92	CGCTGCTGTTGTTTCTGCAATGATATTGGGACTATTACAATCAAGATCTGCGCTG	151		
Qy	128	TGATCAAGTGTGTTTAACTAGTCATGAGAACATATGGCTCATACGCGGAAAGCCATCAT	187		
Db	152	TGATCAAGTGTGTTTAACTAGTCATGAGAACATATGGCTCATACGCGGAAAGCCATCAT	211		
Qy	188	CGTACCGAATGGTGGAGGATCCCGAGAGACCTCCAGTGTACCCCGAGGCGCCAGATG	247		
Db	212	CATATCCCATGGTATCAGAAATCCCGGAGAACCTCGGGTGTGCGTTGAGACCCAGAGCT	271		
Qy	248	AAGGGAACGGTATATGAAGCGGCCACCGTGGAGGTGGCGAGTCTGGTCCATCACCTGC	307		
Db	272	CAGGACAGTGTACGAGCTGCGCGTGTGGAGTGGATGTATCTGCTTCCATCACCTGC	331		
Qy	308	AAGTGCAGCTGCCACCCCGAGGACCTTTCTGCTCTGCGTCTTTAAGCACAGCTCCC	367		
Db	332	AAGTGTGCTGCTCGATGCCCGAGGAGCAATTTCTGCTCTGCGTCTTTAAGCACAGCTCCC	391		
Qy	368	TGGCTCCAGCCGACCTTTGATTTACAAACAGAGGAATCGTTTCCATGCGCATTTTGA	427		
Db	392	TGAATGTCAGCCACATTTTGATTTACAAACAGAGGAGTTGTTTCCATGCTCATTTTGA	451		
Qy	428	ACGTGACAGAGACCCAGGACGAGGAATACCTACTCCATATTACAGAGCAACCGCCACT	487		
Db	452	AAATGACAGAAACCCAGCTGGAGATACCTACTTTTATTCAGAGTGAAGCTACCAAT	511		
Qy	488	ACACAGTACTGTTCCAGTGAATGTAAAGATACACAGCTGTATGTGCTAAGAGAGACTT	547		
Db	512	ACACAAATATTGTTTACAGTGAATATAAGAAATACCTGCTTTACACATTAAAGAGACTT	571		
Qy	548	ACTTTAGAGATGGAAACAGAGATGCACTGCTCTGATCTCCGAGGTTTCCGAGC	607		
Db	572	ACTTTAGAAATATGGAAACAGAGATGCACTGCTCTGATCTCCGAGGTTTCCGAGC	631		
Qy	608	CCACTGTGGAGTGGTGTCTGCACTCCCAAGGAAAGCTGTAAAGAGAGAGCCCTG	667		
Db	632	CGATCGTGGATGGTGTCTGCACTCCCAAGGAAAGCTGTAAAGAGAGAGCCCTG	691		
Qy	668	CTGTTGTGAGAAAGGAGAAAGTACTTCAATGATTTGTTGGAAACAGACATCAGATGCT	727		
Db	692	CTGTTGTGAGAAAGGAGAAAGTACTTCAATGATTTGTTGGAAACAGACATCAGATGCT	751		
Qy	728	GTCTGAGAAATGCACTGGGCGGGAATGCAACAGCTGTTTCAATAGATCTAAACAGG	787		
Db	752	GTCCAGAAATGCACTGGGCGGGAATGCAACAGCTGTTTCAATAGATCTAAACAGG	811		
Qy	788	CTCTCTCAGACACACTCCCGCCAGTTATCTCTGAAAGTGGGGAACCTCTGGATCAGT	847		
Db	812	CTCTCTCAGACACACTCCCGCCAGTTATCTCTGAAAGTGGGGAACCTCTGGATCAGT	871		
Qy	848	GTAAGGCCATCTTGTGAACCATGGATTCTGGGCTCACCTGGAGCTGGAAGACAAAGCCC	907		
Db	872	GCAAGCTGTTTGTGAACCATGGATTCTGGGCTCACCTGGAGCTGGAAGACAAAGCCC	931		
Qy	908	TGAGAGGAGGCACTTGTGATGATGATCTTCTCAAAACAGAGCAATGATTCGGA	967		
Db	932	TCAGAGGAGGCACTTGTGATGATGATCTTCTCAAAACAGAGCAATGATTCGGA	991		
Qy	968	TTCTCTTGGGCTTTGTGCTCTCCGTGGGAGGAGACGACCGGATATTACACCTGCTCTT	1027		

Db	992	TTCTGTTTGTCTTTGTATCATCAGTGGCAGAAACGACACCGATACACTTGTCTCT	1051		
Qy	1028	CCTCAAGCACCCAGCGAGTCCAGGTTGGTACCATCTTAGAAAAAGGTTTATAAAG	1087		
Db	1052	CTTCAAGCATCCAGTCAATCAGCTTTGGTTACCATCTAGGAAAGGATTTATAATG	1111		
Qy	1088	CTACGAGTCCCAAGAGAGATGAAATGTAGCCCGTACGAAAGTTCTGCTTCTCAGTCA	1147		
Db	1112	CTACCAATTTCAAGTGAAGATTTGAAATTTGACCAATATGAAAGATTTTGTCTGTCA	1171		
Qy	1148	GGTTTAAAGCGTACCCAGAAATCCGATGACGATGATCTCTCAAGCCTCATTTCCCT	1207		
Db	1172	GGTTTAAAGCGTACCCAGAAATCCGATGATGATCTCTCAAGCCTCATTTCCCT	1231		
Qy	1208	GTGAACAGAGAGGCTGGAGATGGGTACAGCATATCTAAATTTTGGCATATAAGAAC	1267		
Db	1232	GTGAGCAAAAGGGTCTTGATTAACGATACAGCATATCCAAGTTTGGCATATAAGAAC	1291		
Qy	1268	AGCCAGAGAGTACATATTTCTATGCAAAATGATGACCGCCAGTTTCAACAAATGTTCA	1327		
Db	1292	AGCCAGAGAGTATATATTTCCATGCAAAATGATGATGCCCAATTTACCAAAATGTTCA	1351		
Qy	1328	CGCTGAATATAAGAAAGAAACCTCAAGTGTCTAGCAATGCTCAGCCAGCCAGGCTCT	1387		
Db	1352	CGCTGAATATAAGAAAGAAACCTCAAGTGTCTCGAAGAGCATCGCAAGTCAAGGCTCT	1411		
Qy	1388	GTTCTCTGTATGGCTACCGCTACCTCTTGGACTCTGGAAGAGTGTGGGCAAACTCTC	1447		
Db	1412	GTTCTCGGATGGATACCCATTACCATCTTGGACTCTGGAAGAGTGTGGGCAAACTCTC	1471		
Qy	1448	CCAATTGCAAGGAGAAATCCAGAGAGTGTGGAATATAAGAGCTTAAAGAAAGTGT	1507		
Db	1472	CCAATTGCAAGAGAGATCACAGAGAGTCTGGAATATAAGAGCTTAAAGAAAGTGT	1531		
Qy	1508	TTGCGCAGTGGTGTCCAGCAGTACTCTAAATATGATGAGTGGCGGGAAGAGGCTCTGG	1567		
Db	1532	TTGACAGTGGTGTCCAGCAGTACTCTAAATATGATGAGTGGCGGGAAGAGGCTCTGG	1591		
Qy	1568	TCAAACTCTGTGGTACAAATTTCTATGGCAGCTCTGCGAAACCATCTTTTAAACTCAC	1627		
Db	1592	TCAAGTGTGTGCATACAAATTTCTTGGCAGCTCTGCGAAACCATCTTTTAAACTCAC	1651		
Qy	1628	CAGGCCCTTCCCTTTTCAACAGAACATCTCTTCTATGCGACCATTTGGGCTCTGTC	1687		
Db	1652	CAGGCCCTTCCCTTTTCAACAGAACATCTCTTCTATGCGACCATTTGGGCTCTGTC	1711		
Qy	1688	TCCCTTCAATGTTGTTCTCAITGTTGATCTGCGCAAAATACAAAGAGCAATTTAGGT	1747		
Db	1712	TCCTCTTCAITGTTGTTCTCAITGTTGATCTGCGCAAAATACAAAGAGCAATTTAGGT	1771		
Qy	1748	ACGAGAGTCACTGTCAGATGATCCAGTGTACTGGCCCTTGGATTAACAGTACTTCTACG	1807		
Db	1772	ATGAAGCCAGCTACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	1831		
Qy	1808	TTCACTTCAGGAGTATGAATATGACTTAAAGTGGAGTTCCCGAGAGAGAACTTAGAGT	1867		
Db	1832	TTGATTTCAAGAAATATGAATATGATCTCAATGGGAGTTTCCAGAGAGAAATTTAGAGT	1891		
Qy	1868	TTGGGAGGCTCTGGGCTCTGGGCTTTCCGAGGAGTATGAACCGCCAGCTTATGGCA	1927		
Db	1892	TTGGGAGGCTCTGGGCTCTGGGCTTTCCGAGGAGTATGAACCGCCAGCTTATGGCA	1951		
Qy	1928	TTAGTAAAAAGGAGTCTCAATTTCAAGTGGTGGTGAAGATGCTTAAAGAGAGAGCTGACA	1987		
Db	1952	TTAGTAAAAAGGAGTCTCAATTTCAAGTGGTGGTGAAGATGCTTAAAGAGAGAGCTGACA	2011		
Qy	1988	GCTGTGAAAAAGAGCTCTCATGTCCGAGCTCAAAATGATGACCCACCTGGGACACCATG	2047		
Db	2012	GCTGTGAAAAAGAGCTCTCATGTCCGAGCTCAAAATGATGACCCACCTGGGACACCATG	2071		
Qy	2048	ACAACTCTGATCTGCTGGGGGATGACACTGTTCAGGCGGCACTGCTGATTTTTTGT	2107		

2072	AGAAATATTGTGAACCTGCTGGGGCGTGCACATGTCAGAACCAATTTACTTGATTTTTCG	2133
2108	AAATATTGTTGCTATGTGTGACCTCTCTCAACTACCTCAAGAAGTAAAGAGAGAAGTTTCACA	2167
2132	AATACTGTTGCTATGTGTGATCTTCTCAACTATCTTAAGAGTAAAGAGAGAAAATTTTCACA	2191
2168	GGACATGGACAGAGATTTTTTAAGGAACATAAATTTACAGTTCTTACCCCTACTTTCCAGGCAC	2227
2192	GGACTTTGGACAGAGATTTTCAAGGAACACAAATTTACGTTTTTATCCCCACATTTTCCAATCAC	2251
2228	ATTCAAATTCACGATCGCTCGTTTCACGAGAAAGTTTCAGTTTACACCCGCCCTTGGATCAGC	2287
2252	ATCCAAATTCACGATCGCTCGTTTCACGAGAAAGTTTCAGATACACCCCGACTCGGATCAAA	2311
2288	TCTCAGGGTTCAATGGGAATTCAAATTCATCTGAAAGATGAGATGGAATGAAATGAAAAACGAG	2347
2312	TCTCAGGGCTTCATGGGAATTCATTTTCACTCTGAAGATGAAATTTGAATATGAAAAACCAA	2371
2348	AGAGCTGGCAGAGAAGAGAGAGAGAAATTTGAACGTGCTGACGTTTTGAAGACCTCTCTTT	2407
2372	AAAGGCTG-----GAAGAAAGAGGAGACTTGAATGTGTCTTACATTTTGAAGATCTCTCTTT	2425
2408	GCCTTGGTACCAAGTCGCGCAAAAGGCATGGAAATTCCTGGAGTTTCAAGTCTGCTGTCTCCACA	2467
2426	GCCTTTGCATATCAAGTTTGCCAAAGGAATGGAAATTTCTGGAAATTTAAGTCTGCTGTCTCA	2485
2468	GAGACCTGGCAGCCAGCAATGTGTGTGTCTACCCACGGGAAGGTGGTGAAGATCTGTGACT	2527
2486	GAGACCTGGCGCCAGGAACTGTCTGTCTCACCCACGGGAAGTGGTGAAGATATGTGACT	2545
2528	TTGGACTGGCCCCGAGACATCTGTAGCGACTTCCAGCTACGTGTCAGGGGCAACGCACGGC	2587
2546	TTGGATTGGCTCGAGATATCATGATGATTCCAAATCTGTGTCTCAGGGGCAATCCCGCTC	2605
2588	TGCGGTGAAGTGGATGGCACCCGAGAGCTTATTTGAAGGATCTACACAAATCAAGAGTG	2647
2606	TGCCTGTAAAATGGATGGCCCCGAAAAGCCTGTGTGAAGGCATCTACACCAATTAAGAGTG	2665
2648	ACGTCCTGCTACGGCATCTCTCTGGGAGATATTTTCACTCGGTGTGAACCTTTACC	2707
2666	ATGCTCTGCTCATATGGAAATATTACTGTGGGAATCTTCTCACTTGGTGTGNAATCTTTACC	2725
2708	CTGSCATTCTCTGTCGACGCTAACTTCTATAAATGATTCAGATGGATTTAAAAATGAGC	2767
2726	CTGSCATTCCTGGTTGATGCTAACTTCTACAACTGATTCAAAATGGATTTAAAAATGGATC	2785
2768	AGCCATCTATGCCACAGAGGGATATATCTTGTAAATGCATCTCTGCTGGCTTTTTCAGT	2827
2786	AGCCATTTTATGCTACAGAGAAATATACATTAATAGCAATCTCTGCTGGCTTTTTCAGT	2845
2828	CAAGGAAGCGGCCATCTTCTCCCAACCTGACTTCAATTTTTAGGATGTGAGTGGCAGAGG	2887
2846	CAAGGAACCGGCCATCTTCTCCCTAAATTGACTTCGTTTTTAGGATGTGAGTGGCAGATG	2905
2888	CAGAAGAAG-----CATGTATACAGAACATCCCATCCATCTACC	2924
2906	CAGAAGAAGCGATCTATCAGAATGTGGATGGCCGTGTTTCGGAATGCTTCAACCTTACC	2965
2925	AAAACAGCGCGCCCTCAGCAGAGAGGGCGGCTCAGAGC--CCAGTGGCACAGCGCCAGG	2983
2966	AAAACAGCGCACCTTTTCAGCAGAGAGATGGAAATTTGGGGCTACTCTCTCGCAGGCTCAGG	3025
2984	T-GAAGATTACACAGAGAAAGATTAGCGAGGCGCTTGGACCCCGCC-----ACCTTAG	3038
3026	TCGAAGATTTCGTAGAGGAACAAATTTAGTTTTTAAGGACCTTCATCCCTCCACCTATCCCTAA	3085
3039	CAGGCTGTAGACCGCAGAGCCAAAGATTAGCCTTCGCTCT--GAGGAAGCGCCCTACAGCG	3096
3086	CAGGCTGTAGATTACCAAAAACAAGATAATTTTCATCACTAAAGAAAATCTATTATCAAC	3145
3097	CGTTGCTTCGCTGGACATTTTCTCTAGATGTGCTGTGCCATTACTC-----CAAAGTGA	3149
3146	TGCTGCTTACACAGACTTTTCTCTAGAGCGGCTGTGGCTTTTACTCTGTTTTTCAAAGGGA	3205

[illegible]

[illegible]

Qy	2984	T-GAA	GATTCA	CAGAGAA	AGAGTTAG	CGAGGAG	CGCTTGG	ACCCGCC	-----	ACCCTAG	3038		
Db	3026	TCGA	AGATTTCG	TAGAGGA	CAATTTAG	TTTTTA	AGAGACTT	CATCCCT	CCACCTAT	CCCTAA	3085		
Qy	3039	CAGG	CTGTAG	ACCGCAG	AGCCAA	GATTAG	CGCTCG	CTCT--	GAGGA	AGCGCC	CTACAGCG	3096	
Db	3086	CAGC	TGTAG	ATTAC	CAAAA	CAAGATTA	ATTTTC	ATCA	CTAA	AGAAATCT	ATTATCAAC	3145	
Qy	3097	CGT	TGCTTC	GCTGG	ACTTTTT	CTCTAG	ATGCTGT	CGCATT	TACTC	-----	CAAAGTGA	3149	
Db	3146	TGCT	GCCTT	CAC	CAGACTTTT	CTCTAG	AAGCGCT	CTCG	TTTTTACT	CTTGT	TTTTTCAAAGGGA	3205	
Qy	3150	CTT	CTATA	AAATCA	AACCTCT	CTCTCG	CACAGG	CGGAG	AGCCAA	TATAG	ACTTGT	3209	
Db	3206	CTTT	TGTAA	AAATCA	AAATCAT	CTCTGT	CAAC	AGG	CAGGAG	AGCTG	ATAATGA	3265	
Qy	3210	TGAG	CCCGC	CTAC	CCCTCG	GGGCGCTTT	CCAC	GAGCTT	GAGGG	GAAGCC	ATGATCTGAAA	3269	
Db	3266	GAGC	ATTGAT	CTGC	ATCCA	AGCCCTT	CTC	AG	CGCGCTT	GAGTGA	ATTGTGTACCTGAAG	3325	
Qy	3270	TAT	AGTATAT	TCTT	GTAAAT	ACGTGA	AAACAA	ACCA	ACCGCTTT	TTTTTGT	CTAAGGGAAGC	3329	
Db	3326	TAC	AGTATAT	TCTT	GTAAAT	TACATA	AAACAA	-----	AGCAT	TTTTTGT	CTAAGGGAAGC	3378	
Qy	3330	TAA	ATATG	ATTTT	AAAAAT	CTATG	TTTTT	AAAA	TA	CTATG	TAACTTTT	3389	
Db	3379	TAAT	ATGAT	TTTTT	--	AAGT	CTATG	TTTTT	AAAA	ATAATAT	GTAAATTTTT	3436	
Qy	3390	TGAT	ATATTT	TATG	ATGGA	ATAA	AACTTT	CTACT	GTAA	AAAA	AAAA	3449	
Db	3437	TGAT	ATATTT	TATG	GGTGG	GAATAA	AAATTT	CTCT	ACAG	AAAA	AAAA	3496	
Qy	3450	AAAA	3453										
Db	3497	AAAA	3500										
RESULT 14													
LOCUS	I40109					3476 bp	DNA	linear		PAT 13-MAY-1997			
DEFINITION	Sequence 1 from patent US 5618709.												
ACCESSION	I40109												
VERSION	I40109.1 GI:2083114												
KEYWORDS	Unknown.												
SOURCE	Unknown.												
ORGANISM	Unclassified.												
REFERENCE	1 (bases 1 to 3476)												
AUTHORS	Gewirtz,A.M., Small,D. and Civin,C.I.												
TITLE	Antisense oligonucleotides specific for STK-1 and method for inhibiting expression of the STK-1 protein												
JOURNAL	Patent: US 5618709-A 1 08-APR-1997;												
FEATURES	Location/Qualifiers												
source	1..3476												
BASE COUNT	1044 a	709 c	783 g	940 t									
ORIGIN													
Query Match 63.9%; Score 2204.8; DB 6; Length 3476;													
Best Local Similarity 79.5%; Pred. No. 0;													
Matches 2752; Conservative 0; Mismatches 652; Indels 56; Gaps 10;													
Qy	8	GGCT	ACCG	CGCT	CCGG	AGGCC	ATCG	GGGCG	TGTGG	CGCAG	CGCAG	CGCGGCGTGC	67
Db	35	GGGG	ACCC	CGGCT	CCGG	AGGCC	ATCG	GGGCG	TGTGG	CGCAG	CGCAG	CGCGGCGTGC	91
Qy	68	TGCT	GCTT	GTTT	GTG	TCTC	AGTAA	TGAT	TCTT	GGAC	CCGTT	TACAAA	127
Db	92	CGCT	GCTG	TGTT	TTTT	CTG	CAAT	GATAT	TGGG	AGCT	TATT	ACAAT	151
Qy	128	TGAT	CAAG	TGTG	TTTT	TAA	TCTAG	TCA	TAG	AACA	ATGG	CTCAT	187
Db	152	TGAT	CAAG	TGTG	TTTT	TAA	TCAAT	CA	TAAGA	ACAAT	GATT	TCAT	211

Qy	188	CGTACCGAATGTCGAGGATCCCCAGAAAGACCTCCAGTGTACCCCGAGGCCAGAGTG	247
Db	212	CATATCCCATGGTATCAGATCCCGGGAAGACCTCGGGTGTGGTTGAGACCCAGAGCT	271
Qy	248	AAGGACGCTATATAGAAAGCGCCACCGTGGAGGTGGCCGAGTCTGGGGTCCATCACCCCTGC	307
Db	272	CAGGACAGTGTACGAAGCTGCCGCTGTGGAAGTGGATGTATCTGCTTCCATCACACTGC	331
Qy	308	AAGTGCAGCTGCCACCCAGGGGACCTTTCTGCTGCTCTGGTCTTTAAGCACAGCTCCC	367
Db	332	AAGTCTGCTGCATGCCCGCCAGGGAACATTTCTGCTCTGGGCTTTAAGCACAGCTCCC	391
Qy	368	TGGCTGCCAGCCGACCTTTGATTTACAAAACAGAGGAATCGTTTCCATGGCCATCTCA	427
Db	392	TGAATTGCACGCCACATTTTGAATTTACAAAACAGAGGAGTTGTTTCCATGGTCAATTGA	451
Qy	428	ACGTGACAGAGACAGGAGGAGAAATACCTACTCATATTTACAGAGCGGAACCGCAACT	487
Db	452	AAATGACAGAAACCCAGCTGGAGNATACCTACTTTTATTCAGAGTGAAGCTACCAATT	511
Qy	488	ACACAGTACTGTTCACTGATGTAAGAGATACACAGCTGTATGTGTAAAGAGACCTTT	547
Db	512	ACACAATATTGTTTACAGTGAGTATAAGAAATACCTGCTTTACACATTTAAGAAAGACCTT	571
Qy	548	ACTTTAGAAAGATGGAACCCAGGATGCATGCTCTGCTCATCTCCGAGGGTGTTCGGAGC	607
Db	572	ACTTTAGAAAATGGAACCCAGGAGCGCCCTGGTCTGCATATCTGAGAGCGTTCCAGAGC	631
Qy	608	CCACTGTGAGTGGTGTCTGTCAGCTCCCAAGGAAAGCTGTAAAGAAAGAGCGCCCTG	667
Db	632	CGATCGTGAATGGGTGCTTTGCGATTACAGGGGGGAAGCTGTAAAGAAAGTCCAG	691
Qy	668	CTGTTGTGAGAAAGAGGAAAGGTTCTTCATGAGTGTTCGGAACAGACATCAGATGCT	727
Db	692	CTGTTGTTTAAAGAGGAGGAAAGTCTTCATGAATTTATTTGGGACGACATAAGGTGCT	751
Qy	728	GTGCTAGAAAATGCACTGGGCGCGGAATGCAACCAAGCTGTTTACCATAGATCTAAACAGG	787
Db	752	GTGCCAGAAATGAATGGGCGAGGGAATGCACAGAGCTGTTTCAATAGATCTAAATCAA	811
Qy	788	CTCCTCAGAGCACACTGCCCCAGTTTATTCCTCAAGTGGGGGAACCTTTGTGATCAGGT	847
Db	812	CTCCTCAGACCAATGCCACATTTATTTCTTAAAGTAGGGGAACCTTTATGATAGGT	871
Qy	848	GTAGCGCCATCCATGTGAACCAATGGATTGCGGCTCACCTGGGAGCTGGAAGACAAAGCCC	907
Db	872	GCAAGCTGTTTCAATGTGAACCATGGATTTCGGGCTCACCTGGGAATTTAGAAAAACAAAGCAC	931
Qy	908	TGGAGAGGGCAGCTACTTTTGATGAGTACTTCTTCAAAACAGGACCATGATTCGGA	967
Db	932	TCGAGGAGGGCAACTACTTTTGATGATGATGATCTTCAACAAACAGACATATGATACGA	991
Qy	968	TTCTCTTGGCTTCTGCTTCCGTTGGGAGGAACGACACCGGATATTACCTGCTCTT	1027
Db	992	TTCTGTTGCTTTTGTATCATCAGTGGCAAGAAACGACACCGGATACTACACTTGTCT	1051
Qy	1028	CCTCAAAAGCACCCAGCCAGTCAAGCTTGGTGACCATCTCTAGAAAAGGGTTTATAACG	1087
Db	1052	CTTCAAGCATCCAGTCAATCAGCTTTGTTGTTTACCATCGTAGAAAAGGGATTTATAATG	1111
Qy	1088	CTACAGCTCGCAAGAAGATGAAATTCACCGGTGACCGGTGCGAAGTTCTGCTCTCAGTCA	1147
Db	1112	CTACCAATTCAAAGTGAAGATTATGAAATTCACCAATATGAAAGTGTGTTTTCTGTCA	1171
Qy	1148	GGTTTAAAGCTGACCCAGAAATCCGATGCACTGATCTTCTCTCAAGCTCATTTCCCT	1207
Db	1172	GGTTTAAAGCTTACCCACAAATCAGATGTACGGGACCTTCTCTCGAAAATCATTTCCCT	1231
Qy	1208	GTGACAGAGAGCGCTGGAGGATGGGTACAGATATCTAAATTTTGCATATAAGAAC	1267
Db	1232	GTGACAAAGGGTCTTTGATAACGGATACAGCAATTCGAATTTTGAATCATAGCAC	1291

Db 3437 TGATATATTTTATGGTGGGAATAAAATTTCTACTACAGA 3476

RESULT 15
AX695522
LOCUS 3475 bp DNA linear PAT 31-MAR-2003
DEFINITION Sequence 1149 from Patent WO03008583.
ACCESSION AX695522
VERSION AX695522.1 GI:29418674
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE
AUTHORS Morris, D.W. and Engelhard, E.K.
TITLE Novel compositions and methods for cancer
JOURNAL Patent: WO 03008583-A 1149 30-JAN-2003;
Sagres Discovery (US)
FEATURES
Location/Qualifiers
1..3475
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 1042 a 709 c 784 g 940 t
ORIGIN

Query Match 63.8%; Score 2203; DB 6; Length 3475;
Best Local Similarity 79.6%; Pred. No. 0;
Matches 2749; Conservative 0; Mismatches 650; Indels 56; Gaps 10;

QY 8 GGCTACCGCGCTCGGAGGCCATCGGCGGTGGCGAGCGACGCGCGCGCTGC 67
DB 35 GGGACCCCGGCTCGGAGGCCATCGGCGGTGGC---GGCGAGCGGGCACGCTGC 91
QY 68 TGCTGCTGTTGTTGTCAGTAATGATTTCTTGAGACCGTTACAAACCAAGACCTGCTG 127
DB 92 CGCTGCTGCTGTTTCTGCAATGATATTTGGGACTATTACAAATCAAGATCTGCTG 151
QY 128 TCATCAAGTGCTTTTAAATCAGTCATGAGAACATGGCTCATCAGCGGGAAGCCATCAT 187
DB 152 TGATCAAGTGCTTTTAAATCAATCATAGAACAAATGATCATCATGTTGGGAGTCAATCAT 211
QY 188 CGTACCGAATGTGCGAGGATCCCAAGAGACCTCCAGTGTATCCCGAGCGCGCAGAGTG 247
DB 212 CATATCCATGTTATCAGAAATCCCGGAGACCTCGGGTGTGGTTGAGACCCAGAGCT 271
QY 248 AAGGACGGTATATGAGCGGCCACCGTGAGGTGCGGAGTCTGGGTCCATCACCGCTGC 307
DB 272 CAGGGACAGTATGAGGAGCTGCGGCTGTGGAAGTGGATGTATCTGCTTCCATCACACTGC 331
QY 308 AGTGCAGCTCGCCACCCAGGGGACCTTTCTGCTCTGCGCTCTGGGTCCTTAAGCAGCTCCC 367
DB 332 AAGTGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 391
QY 368 TGGGCTGCCAGCGCACTTTGATTTTACAAACAGAGGAATCGTTTCCATGGCCATCTTGA 427
DB 392 TGAATGCCAGCACATTTTGAATTTTACAAACAGAGGATGTTTTCATGGTCAATTTGA 451
QY 428 ACGTGACAGAGACCCAGGAGGAATATCCTACTCCATATTCAGAGCGAACCGCGCAACT 487
DB 452 AAATGACAGAAACCCCAAGCTGGAGATACCTACTTTTATTCAGAGTGAAGCTACCAAT 511
QY 488 ACACAGTACTGTTTACAGTGAATGTAAGATACACAGCTGTATGTCTAAGGAGACCTT 547
DB 512 ACACAAATATTTTACAGTGAATGTAAGAAATACCTGCTTTTACACATTAAGAGACCTT 571
QY 548 ACTTTAGGAAGTGAAGAACCCAGAGTGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 607
DB 572 ACTTTAGAAATATGAAGAACCCAGGACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 631
QY 608 CCACTGTGAGTGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 667

Db 632 CGATCGTGAATGGGTGCTTTTGGATTTACAGGGGAAAGCTCTAAAGAAAGAAAGTCCAG 691
QY 668 CTGTTGTCAGAAAGAGGAAAGGTACTTCTCATGAGTTGTTGCGAAACAGACATCAGATGCT 727
DB 692 CTGTTGTTTAAAGAGGAGGAAAGTCTTTCATGAATTTATTTGGGACGGACATAGGTGCT 751
QY 728 GTGCTAGAAATCAGCTGGGCGCGGAATGACCAAGCTGTTTACCATAGATCTTAAACACAG 787
DB 752 GTGCCAGAAATGAATGCGGCGAGGAATGACCAAGCTGTTTCAATAGATCTTAAATCAA 811
QY 788 CTCCTCAGAGCACATGCCCCCAGTTTATCTGAAAGTGGGGGAAACCTTGTGGATCAGGT 847
DB 812 CTCCTCAGACCAATTTGCCCCAATTTATTTCTTAAAGTAGGGGAACCTTATGATAGGT 871
QY 848 GTAAGCCATCCATGTGAACCATGCGGCTCACCTGGGAGCTGGAAGACAAAGCCC 907
DB 872 GCAAGCTGTTTATGTGAACCATGGAATCGGGCTCACCTGGGAATTAGAAAAAACAAGCAC 931
QY 908 TGGAGGAGGCGAGCTACTTTTGATGAGTACCTACTCCCAAAACAGACCATGATTCGGA 967
DB 932 TCGAGGAGGCGAATCTTTTGAATGAGTACCTTATTCACAAACAGAACTATGATACGGA 991
QY 968 TTCTCTTGGCTTTTGTGCTTTCGTTGGGAAGGAACGACACCGGATATTACACCTGCTCTT 1027
DB 992 TTCTGTTTGTGTTTGTATCATCAGTGGCAAGAAACGACACCGGATCTACACTGTTCT 1051
QY 1028 CTTCAAAGCACCCAGCAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1087
DB 1052 CTTCAAAGCATCCAGTCAATCAGCTTTGTTTACCATCGTAGGAAGGATTTATAAATG 1111
QY 1088 CTACGAGCTCGCAAGAGAGTATGAAATTTGACCCGTACGAAAGATTTGCTTCTCAGTCA 1147
DB 1112 CTACCAATTTCAAGTGAAGATTTGAAATTTGACCAATATGAAGATTTTGTGTTCTGTC 1171
QY 1148 GGTTTAAAGCTACCCACCAATCCGATGCACTGGGATCTTCTCTCAAGCTCATTTTCTT 1207
DB 1172 GGTTTAAAGCTACCCACCAATCCGATGCACTGGGATCTTCTCTCAAGCTCATTTTCTT 1231
QY 1208 GTGAAACAGAGAGCGCTGGAGGATGGGTACAGCATATCTAAATTTTGCATCATAAAGA 1267
DB 1232 GTGAGCAAAAGGCTCTTGATACGGATACAGCATATCCAAATTTGCAATCATAAAGCAC 1291
QY 1268 AGCAGGAGATGATATTTCTATGCAAGAAATGATGACGCCAGCTTCCACCAATGTTCA 1327
DB 1292 AGCAGGAGATATATTTCCATGCAAGAAATGATGATGATGATGATGATGATGATGATG 1351
QY 1328 CGCTGAATATAGAAAGAACCTTCAAGTGTCTAGCAATGCTCAGCGAGCCAGCGCTCT 1387
DB 1352 CGCTGAATATAGAAAGAACCTTCAAGTGTCTAGCAATGCTCAGCGAGCATCGGCGCT 1411
QY 1388 GTTCTCTGATGGCTACCGCTACCGCTTGGACCTGGAGAAAGTGTTCGGAACAAATCTC 1447
DB 1412 GTTCTCGATGATACCTTACCATCTTGGACCTGGAGAAAGTGTTCAGACAGTCTC 1471
QY 1448 CCAATTTGCAAGGAGAAATCCCAAGAGGATTTGGAATATAAAGGCTTAAACAGAAAGT 1507
DB 1472 CCAACTGCAAGAGAGATCAAGAGAGGAGTCTGGAATAGAAAGGCTTAAACAGAAAGT 1531
QY 1508 TTGGCCAGTGGGTGCTGAGCAGTACTTAAATATGATGATGATGATGATGATGATGATG 1567
DB 1532 TTGGCAGTGGGTGCTGAGCAGTACTTAAATATGATGATGATGATGATGATGATGATG 1591
QY 1568 TCAATGTCTGCTGATACAAATCTTATGGGACGCTTTGCGAAACCATCTTTTAAACTCAC 1627
DB 1592 TCAAGTGTCTGATACAAATCTTCTTGGACATCTTGTGAGACGATCTTTTAAACTCTC 1651
QY 1628 CAGGCCCTTCCCTTTTCAAGACAAATCTTCTTATGCGCAATTTGGGCTCTGTC 1687
DB 1652 CAGGCCCTTCCCTTTTCAAGACAAATCTTCTTATGCAACAAATTTGGTGTGTTGTC 1711
QY 1688 TCCCTTCAATTTGTTGTTTCAATTTGTTGATCTGCGCAAAATACAAAGCAATTTAGGT 1747
DB 1712 TCCTCTTCAATTTGTTGTTTCAATTTGTTGATCTGCGCAAAATACAAAGCAATTTAGGT 1771

QY	1748	ACGAGAGTCAGCTGCAGATGATCCAGGTGACTGGCCCTCGATAACGAGTACTTCTACG	1807
DB	1772	ATGAAGCCAGCTACAGATGGTACAGGTGACGGCTCTCTCAGATAATGAGTACTTCTACG	1831
QY	1808	TTGACTTCAGGGACTATGAATATGACCTTAAAGTGGGAGTTCGCCGAGAGAACTTAGAGT	1867
DB	1832	TTGATTTTCAGAGAAATAGAAATATGATCTCAATGGGAGTTTCCAAGAGAGAAATTTAGAGT	1891
QY	1868	TTGGGAAGGTCCTGGGGTCTCGCGCTTTCGGGAGGGTGATGAACGCCACGGCCCTATGGCA	1927
DB	1892	TTGGGAAGGTTACTAGGATCAGGTGCTTTTGGAAAAGTGATGAACGCCACAGCTTATGGAA	1951
QY	1928	TTAGTAAACGGGAGTCTCAATTCAGGTGGCGGTGAGATGCTAAAGAGAGAAAGCTGACA	1987
DB	1952	TTAGCAAAACAGGAGTCTCAATTCAGGTTCGGTCCGCTCAAATGCTGAAAGAAAAGCAGACA	2011
QY	1988	GCTGTGAAAAAGAAAGCTCTCATGTGCGAGCTCAAAAATGATGACCCACCTCGGGAACCCATG	2047
DB	2012	GCTCTGAAAGAGAGGCACTCATGTCAAGAACTCAAGATGATGACCCAGCTGGGAAGCCACG	2071
QY	2048	ACAACATCGTGAATCTGCTGGGGGATGCACACTGTCAAGGCCAGTGATCTTGATTTTTTG	2107
DB	2072	AGAAATATTTGTAACCTGCTGGGGCGTGCACTGTCAAGACCAATTTACTTGATTTTTTG	2131
QY	2108	AATATTGTGCTATGGTGACCTCTCAACTACCTAAGAACTAAAGAGAGAGAGTTTCACA	2167
DB	2132	AATACTGTGTATGGTGATCTTCTCAACTATCTAAGAGTAAAGAGAGAAAAATTTTCACA	2191
QY	2168	GGACATGGACAGAGATTTTTAAGGAAACATAATTTTCAGTTCTTACCTTACTTTCCAGGCAC	2227
DB	2192	GGACTTGGACAGAGATTTCAAGGAAACACAATTTTCAGTTTATCCCACTTCCAATCAC	2251
QY	2228	ATTCAAAATTCAGCATGCGTGGTTTCAGAGAGTTTCAGTTACACCCGCCCTTGGATCAGC	2287
DB	2252	ATCCAAAATTCAGCATGCGTGGTTTCAGAGAGAGTTTCAGATACACCCGGACTCGGATCAA	2311
QY	2288	TCTCAGGGTTCAATGGGAATTCAAATTCATCTGAAGATGAGTTGAATATGMAAACGAGA	2347
DB	2312	TCCTCAGGGCTTCATGGGAATTCATTTCACTCTGAAGATGAAATGGAATATGMAAACCAAA	2371
QY	2348	AGAGGCTGGCAGAAAGAGAGGAGGAAGATTGAAACGTGCTGACGTTTGAAGACCTCCCTT	2407
DB	2372	AAAGGCTG-----GAAGAGAGAGGAGACTTGAATGTGCTTACATTTGAAGATCTTCTTT	2425
QY	2408	GCCTTGGTACCAGTGGCCAAAGGCATGGAATCTCTGGAGTTCAAGTCTGTGTGCCACA	2467
DB	2426	GCCTTGCATATCAAGTTTGCCAAAAGGAATGGAATTTCTGGAAATTTAAGTCTGTGTGTTACA	2485
QY	2468	GAGACCTGGCAGCCAGGAATGTGTGTGTACCCACGGGAAGGTGGTGAAGATCTGTGACT	2527
DB	2486	GAGACCTGGCCGCCAGGAACTGCTTGTCAACCACGGGAAAGTGGTGAAGATATGTGACT	2545
QY	2528	TTGGACATGGCCCGAGACATCCTGAGGACATCCAGCTACGTCGTGAGGGCAACGCAAGGC	2587
DB	2546	TTGGATTTGGCTCGAGATATCATGAGTGATTCACACTATGTTGTGAGGGCAATGCCCGTC	2605
QY	2588	TGCCGCTGAAGTCGATGGCAACCGAGAGCTTATTTGAAGGATCTACAACTCAAGAGTG	2647
DB	2606	TGCCCTGTAAATGGATGGCCCCGAAAGCCTGTTTGAAGGCATCTACCATTTAAGAGTG	2665
QY	2648	ACGCTCGGTCTACGGCATCCTTCTCTGGGAGATATTTTCACTGGGTGTGAACCCCTTACC	2707
DB	2666	ATGCTCGTCATATGGAATATTACTGTGGGAATCTTCTCACTTGGTGTGAATCCCTTACC	2725
QY	2708	CTGGCATTTCTGTGCGCTAACTTCTTATAAAGTGAATCAGAGTGAATTTAAAGTGAAGC	2767
DB	2726	CTGGCATTCGGGTGTGCTAACTTCTCAAACTGATTCAAAATGGAATTTAAAAATGGAATC	2785
QY	2768	AGCCATCTCATGCCACAGAGGATATATCTTTGTAATGCAATCCCTGCTGGCTTTTGACT	2827
DB	2786	AGCCATTTTATGCTACAGAGAAATATACATTAATGAATTCGAATCTGCTGGCTTTTGACT	2845

Search completed: August 28, 2003, 01:02:04
Job time : 16115.1 secs

Qy	2828	CAAGGAAGCGGCATCTTCCCCAAACCTGACTCAATTTTATAGATGTCAGCTGGCAGAGG	2887
Db	2846	CAAGGAACAGGCCATCTCTCCCTAAATTGACTTCGTTTTTATAGATGTCAGCTGGCAGATG	2905
Qy	2888	CAGAAGAAG-----CATGTATCAGAAATCCATCCATCTACC	2924
Db	2906	CAGAAGAAGCGATGTATCAGAAATGTGGATGGCGTGTTTCGGAATGTCTTCACACTTACC	2965
Qy	2925	AAAAAGCGGCGCCCTCAGCAGAGAGCGGGCTCAGAGC-CCAGTGCCCACACGCGCAGG	2983
Db	2966	AAAAAGCGGACCTTTTCAGCAGAGAGATGGATTTGGGGCTACTCTCTCCGCAAGCTCAGG	3025
Qy	2984	T-GAAGATTACAGAGAGAAGAAGTTACGAGAGAGCGCTTGGACCC-----CGCCACCCCTAG	3038
Db	3026	TCGAAGAATTGCTAGAGGAACAAATTATAGTTTAAAGGACTTCATCCCTCACTATCCCTAA	3085
Qy	3039	CAGGCTGTAGACCGCAGAGCAGAGATTAGCCTCGCCTCT--GAGGAAGCGCCCTCAGCG	3096
Db	3086	CAGGCTGTAGATTACCAAACAAGATTAAATTTTCATCACTAAAGAANAATCTATTATCAAC	3145
Qy	3097	CGTTGCTTCGCTGGACTTTTCTCTAGATGCTGTGCCATTACTC-----CAAAGTGA	3149
Db	3146	TGCTGTCTTACCAGACTTTTCTCTAGAAGCGCTCTCGTTTACTCTTGTTTTCAAAGGGA	3205
Qy	3150	CTTCTATAAAATCAAACCTCTCTCCGACAGGCGGAGAGCCAAATATGAGACTTGTTCG	3209
Db	3206	CTTTTGTAAATCAAATCATCTGTGTCAAGGCAGGAGAGCTGATAATGAACCTTTATTG	3265
Qy	3210	TGAGCCGCGCTACCCCTGGGGCGCTTTCCACGAGCTTGTAGGGGAAAGCCATGTATCTGAAA	3269
Db	3266	GAGCATTTGATCTGCATCCCAAGGCTTCTCAGGCCGCTTGTAGTGAAATTTGTACTTGAAG	3325
Qy	3270	TATAGTATATCTTGTGAAATACTGTGAAACAAACCAAAACCCGTTTTTGTCTAAGGGAAAGC	3329
Db	3326	TAACAGTATATCTTGTGAAATACATAAAACAAA-----AGCATTTTGTCTAAGGGAAGC	3378
Qy	3330	TAAATATGATTTTAAAAATCTATGTTTTTAAAAATCACTATGTAACTTTTTCATCTATTTAG	3389
Db	3379	TAATATGATTTTTT--AAGTCTATGTTTTTAAATAATATGTAAATTTTTTCAGCTATTAG	3436
Qy	3390	TGATATATTTTATPGGATGAAATATAACTTCTTACT	3424
Db	3437	TGATATATTTTATPGGTGGGAATAAAATTTCTACT	3471

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 27, 2003, 15:17:42 ; Search time 8334.07 Seconds
(without alignments)
10069.907 Million cell updates/sec

Title: US-09-919-408A-1
Perfect score: 3453
Sequence: 1 CGCGCTGCTACCGCGCGC.....AAAAAAAAAAAAAAAAAAAA 3453

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:

1: em_estba:*

2: em_esthum:*

3: em_estin:*

4: em_estmu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_htc:*

9: gb_est1:*

10: gb_est2:*

11: gb_hic:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_estom:*

17: em_gss_hum:*

18: em_gss_inv:*

19: em_gss_pin:*

20: em_gss_vrt:*

21: em_gss_fun:*

22: em_gss_mam:*

23: em_gss_mus:*

24: em_gss_pro:*

25: em_gss_rod:*

26: em_gss_pig:*

27: em_gss_vxl:*

28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2080.8	60.3	3647	11	BC036028 Homo sapi
2	1229.8	35.6	3761	11	AK045865 Mus muscu
3	655.8	19.0	765	9	AI323253
4	556.8	16.1	618	9	AA120050

5	531.8	15.4	604	10	BF523018
6	502.4	14.5	747	9	AI323643
7	495.4	14.3	500	13	BU696796
8	486.6	14.1	522	10	BB644407
9	452.6	13.1	621	12	BI360262
10	446.6	12.9	658	14	BY727686
11	445	12.9	445	10	BE651447
12	432.8	12.5	714	12	BI461248
13	428.4	12.4	597	13	BU695104
14	426.2	12.3	515	13	BO556177
15	361.6	10.5	403	12	BM117950
16	354.4	10.3	507	12	BM484050
17	340.8	9.9	460	9	AV713950
18	333.6	9.7	3312	11	AK004947
19	333.6	9.7	3616	11	AK080326
20	333.6	9.7	3687	11	AK079247
21	332	9.6	3822	11	BC050024
22	327.4	9.5	3687	11	AK076215
23	324.8	9.4	3883	11	AK046795
24	324.8	9.4	5174	11	BC026713
25	314	9.1	394	10	BG145808
26	313.8	9.1	3887	11	AK047010
27	298.8	8.7	451	14	CB787035
28	294	8.5	457	9	AA996414
29	279.4	8.1	332	14	CB605696
30	277.6	8.0	357	9	AI556371
31	273	7.9	424	9	AI180146
32	266	7.7	502	10	BF557871
33	265.8	7.7	455	12	BM147250
34	254	7.4	568	10	BE679272
35	253	7.3	318	14	CB548474
36	248.2	7.2	375	9	AA851570
37	245.8	7.1	939	13	BX371699
38	244.8	7.1	660	10	BE047675
39	235.8	6.8	724	14	BY764070
40	234.8	6.8	640	14	BY749257
41	233	6.7	449	13	BY545844
42	227.8	6.6	386	9	AI178876
43	227.6	6.6	407	9	AV596131
44	226.6	6.6	739	10	BF159205
45	226.6	6.6	779	14	CA511669

ALIGNMENTS

RESULT 1
BC036028
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
TITLE
JOURNAL
REMARK
COMMENT

BC036028 3647 bp mRNA linear HTC 23-SEP-2002
Homo sapiens, clone IMAGE:5272266, mRNA.
BC036028
BC036028.1 GI:23273564
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Strausberg, R.
Direct Submission
Submitted (31-JUL-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgaps-r@mail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305

Db 1790 TTTAGGTATGAAGCCAGCTACAGATGGTACAGGTGACCGGCTCTCAGATATAGTAC 1849
Qy 1801 TTCTACGTTGACTTCAGGGACT-ATCAATATACCTTAAAGTGGGAGTTCCCGAGAGAGAA 1859
Db 1850 TTCTACGTTGATTTCAGAGAAATGAATATATCATCTCAATGGGAGTTTCCAGAGAGAAA 1909
Qy 1860 CTTAGAGTTTGGGAAGGTCTGGGCTCTGGGCTTTCCGGAGGGTGTAGAAAGCCACGGC 1919
Db 1910 TTTAGAGTTTGGGAAGGTCTAGGATCAGGTCTCTTTGGAAAGTGTAGAACGCCACAGC 1969
Qy 1920 CTATGGCAATAGTAAACCGGAGTCTCAATTCAGGTGGCGGTGAAGATCTATAAGAGAA 1979
Db 1970 TTATGGAAATAGCAAAACAGGAGTCTCAATCAGGTGGCGGTCAAAATCTGAAGAGAAA 2029
Qy 1980 AGCTCACAGCTGTGAAGAGAGCTCTCATGTGGAGCTCAAAATGATGACCCACCTGGG 2039
Db 2030 AGCAGACAGCTCTGAAGAGGAGCTCATGTGAGACTCAAGATGATGACCCAGCTGGG 2089
Qy 2040 ACACCATGACAACTGCTGAATCTGTGGGGGATGCACTGTGAGGGCCAGTGTACTT 2099
Db 2090 AAGCCACGAGAAATTTGAACTCTGGGGGCTGCACACTGTGAGGACCAATTTACTT 2149
Qy 2100 GATTTTGAATTTGCTGTATGGTGACCTCTCACTACTTAAGAGTAAAGAGAGAA 2159
Db 2150 GATTTTGAATTTGCTGTATGGTGATCTTCTCACTATCTAAGAGTAAAGAGAGAAA 2209
Qy 2160 GTTTCACAGGACATGACAGAGATTTTAAAGAACATAATTTCACTTCACTACTTT 2219
Db 2210 ATTTTACAGGACTTGAACAGAGATTTTCAAGGAAACAAATTTCACTTTTACCCACTTT 2269
Qy 2220 CCAGGCACATTCAAATTCAG----- 2240
Db 2270 CCAATCACATCCAAATTCAGTAAAGAAAGATGAGCTTTTACAAAGGCAAACTGGAAGAAA 2329
Qy 2241 ----- 2240
Db 2330 GAAGGATGGTGAACCGCTTACGGGACTCTCGGAGATCTGTATTATGTGAGGAGAAAGTG 2389
Qy 2241 -----CATGCTGGTTACAGAGAGTTTCAGTT 2267
Db 2390 GGCTGAGCTCAGAAACCAAGAAATGAGATCGATCATGCTGGTTCAAGAGAGTTTCAGAT 2449
Qy 2268 ACACCGGCTTGGATCAGCTCTCAGGGTTCAATGGGAAATCAATTCATTCTGAAGATGA 2327
Db 2450 ACACCGGACTCGGATCAATCTCAGGGCTCATGGGAAATCAATTCATCTGGAAGATGA 2509
Qy 2328 GATTGAATGAAACCAAGAGGCTGGCAGAGAGAGGAGAGAGATTTGAACGTGCT 2387
Db 2510 AATTGAATGAAACCAAGAGGCTG-----GAAGAGAGGAGGACTTGAATGTGCT 2563
Qy 2388 GACGTTTGAAGACCTCTTGTGCTTACCAAGTGCCAGAGGCATGGAATTCCTGGA 2447
Db 2564 TACATTTGAAGATCTTCTTGTCTTGCATATCAAGTTGCAAGGAAATGGAATTCCTGGA 2623
Qy 2448 GTTCAAGTCTGTGTCCACAGAGACCTGGCAGCAGGAATGTGTGCTCACCCACCGGAA 2507
Db 2624 ATTTAAGTCTGTGTTCAAGAGACTGGCGCCAGGAAAGTGTGTGCTCACCCACCGGAA 2683
Qy 2508 GGTGTGAAGATCTGTGATTTGGATGGCCCGAGACATCTGAGGAGCTCCAGTACGT 2567
Db 2684 AGTGTGAAGATGTGACTTTGGATTGGCTCGAGATATCATGATGATTCCAACTATGT 2743
Qy 2568 CGTCAGGGCAACGACGCTGGCTGAGTGGATGGCACCCGAGAGCTTATTGAAGG 2627
Db 2744 TGTCAAGGGCAATGCCCGCTGCTGCTGTAAGATGGATGGCCCCGAAAGCTGTGTAAGG 2803
Qy 2628 GATCTACAAATCAAGAGTGAAGTCTGCTTACGCGATCCTTCTCTGGGAGATTTTTC 2687
Db 2804 CATCTACACATTAAGAGATGTCTGTCATATGGAATATTAATGAGGAAATCTTCTC 2863
Qy 2688 ACTGGGTGAACCCCTTACCTGGCATTCCTGTCGACGCTAACTTCTATAAATGATTCA 2747
Db 2864 ACTTGGTGTGAATCCTTACCTGGCATTCGGTTGATGCTAACTTCTACAAATGATTCA 2923

Qy 2748 GAGTGGATTTAAATGGAGCAGCCATTTATGCCACAGAGGGATATATCTTTTGAATGCA 2807
Db 2924 AAATGGATTTAAATGGATCAGCCATTTATGCTACAGAGAAATATATCATTTATAATGCA 2983
Qy 2808 ATCTCTGCTGGGCTTTTGAATCAAGAGAGCGGCATCTTTCCCAACCTGACTTCATTTT 2867
Db 2984 ATCTCTGCTGGGCTTTTGAATCAAGAGAGCGGCATCTTTCCCTAATTTGACTTCGTTT 3043
Qy 2868 AGGATGTGAGCTGGCAGAGGAGAGAAAG-----CATGTATC 2904
Db 3044 AGGATGTGAGCTGGCAGAGATGCAGAGAGAGCGATGTATCAGAAATGTGATGGCGGTGTTT 3103
Qy 2905 AGAATCATCATCTTATCTACCAAAAAGCGGCGCCCTCAGCAGAGAGCGGGCTCAGAGC- 2963
Db 3104 GGAATGTCTCACACTTACCAAAAAGCGGAGCTTTTTCAGCAGAGAGATGATTTGGGGCT 3163
Qy 2964 CCAGTGGCAGAGCGCCAGGT-GAAGATTCACAGAGAGAGAGATTTAGCGAGAGGCGCTTG 3022
Db 3164 ACTCTCTCGCAGGCTCAGGTGGAAGATTCGTAGAGGAAATTTAGTTTAAAGGACTTC 3223
Qy 3023 GACCC-----CGCCACCTAGCAGGCTGTAGACCGCAGAGCCAGATTAGCCTCGCCTCT- 3077
Db 3224 ATCCCTCCACCTATCCCTAACAGGCTGTAGATTACCAAAACAAGATTAATTTCACTA 3283
Qy 3078 -GAGGAAGCGCCCTACAGCGCTTGTCTCGTGGACTTTTCTCTAGATGCTGTGCTCAT 3136
Db 3284 AAAGAAATCTATTATCAACTCTGCTTCCACAGACTTTTCTCTAGAGGCTGTCTGCTT 3343
Qy 3137 TACTC-----CAAGTGACTTCTATAAAATCAAAACC-TCTCCTGCACAGCGGGAGA 3188
Db 3344 TACTCTGTTTCAAAGGGACTTTTGTAAATCAAAATCAATCTCTGTCAAGCGCAGGAGGA 3403
Qy 3189 GCAATAATGAGACTTGTGGTGAGCGCCCTACCTCGGGGCTTTTCCACGAGCTTGAG 3248
Db 3404 GCTGATATGAATTTATTTGGAGCATTTGATCTGCATCCAGGCTTCTCAGGCTGGCTTG 3463
Qy 3249 GGGAAAGCCATCTATCTGAAATATAGTATATTCTTGTAAATACGTGAAACCAAAACC 3308
Db 3464 AGTGA- TTGTGTACTCTGAAGTACAGTATATCTTGTAAATACATAAAACAAA----- 3515
Qy 3309 CGTTTTCCTAAGGAAAGCTAAATATGATTTTAAATCTATGTTTAAATACTAT 3368
Db 3516 AGCATTTTCTAAGGAGAGCTAAATATGATTTT--AAGTCTATGTTTAAATATATAT 3573
Qy 3369 GTAACTTTTCTATTTAGTGATATATTTATGATGAGATAAATTTTCTACTGTA 3428
Db 3574 GTAAATTTTCTAGCTATTTAGTGATATATTTTATGGTGGAAATAAATTTCTACTCAA 3633
Qy 3429 AAAAAAAAAAAAAA 3442
Db 3634 AAAAAAAAAAAAAA 3647

RESULT 2

AK045865 3761 bp mRNA linear HTC 05-DEC-2002
LOCUS AK045865
DEFINITION Mus musculus adult male corpora quadrigemina cDNA, RIKEN full-length enriched library, clone:B20315G04 product:FMS-like tyrosine-kinase 3, full insert sequence.
ACCESSION AK045865
VERSION AK045865.1 GI:26337656
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1
AUTHORS Carninci, P. and Hayashizaki, Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636


```
QY 481 GCCAACTACAGTACTGTTTCACAGTGAATGAAGATACACAGCTGTATGCTGAAGG 540
Db 498 GCCAACTACAGTACTGTTTCACAGTGAATGAAGATACACAGCTGTATGCTGAAGG 557
QY 541 AGACCTTACTTTAGGAAGATGGAACACCAAGGATGCACTGCTGTGCACTCTCCAGAGGTT 600
Db 558 AGACCTTACTTTAGGAAGATGGAACACCAAGGATGCACTGCTGTGCACTCTCCAGAGGTT 617
QY 601 CCGAGGCCACATGTGGAGTGGTGTCTGCAAGTCCCAAGGAAAGCTGTAAAGAA 660
Db 618 CCAGAGGCCACATGTGGAGTGGTGTCTGCAAGTCCCAAGGAAAGCTGTAAAGAA 677
QY 661 GGCCCTGCTGTTGTCAGAAAGGAGGAAAGTACTTCAATGAGTGTTCGGAACAGATC 720
Db 678 GGCCCTGCTGTTGTCAGAAAGGAGGAAAGTACTTCAATGAGTGTTCGGAACAGATC 737
QY 721 AGATGCTGTGCTAGAAATGCACTGGGCCGCGAATGACCAAGCTGTTCACCATAGATCTA 780
Db 738 AGATGCTGTGCTAGAAATGCACTGGGCCGCGAATGACCAAGCTGTTCACCATAGATCTA 797
QY 781 AACAGGCTCTTCAGAGCACATGCCCCAGTTATTCTGAAAGTGGGGAAACCTTTGTGG 840
Db 798 AACAGGCTCTTCAGAGCACATGCCCCAGTTATTCTGAAAGTGGGGAAACCTTTGTGG 857
QY 841 ATCAGGTGAAGGCCATCCTATGTAACCATGGAATCGGGCTACCTGGGAGCTGGAAGAC 900
Db 858 ATCAGGTGAAGGCCATCCTATGTAACCATGGAATCGGGCTACCTGGGAGCTGGAAGAC 917
QY 901 AAAGCCCTGGAGGAGGAGGAGTACTTTGAGATGAGTACTTCTCCACAAACAGGACCATG 960
Db 918 AAAGCCCTGGAGGAGGAGGAGTACTTTGAGATGAGTACTTCTCCACAAACAGGACCATG 977
QY 961 ATTCGGATTTCTCTGGCCCTTTGTGCTCCGCTGGGAAGAACACACCGATATTACACC 1020
Db 978 ATTCGGATTTCTCTGGCCCTTTGTGCTCCGCTGGGAAGAACACACCGATATTACACC 1037
QY 1021 TGCTTTCTCTAAAGCACCCACGACAGTTCAGCGTTGGTGACCATCTCTCAAGGCTCA 1080
Db 1038 TGCTTTCTCTAAAGCACCCACGACAGTTCAGCGTTGGTGACCATCTCTCAAGGCTCA 1097
QY 1081 ATAAAGCTTACAGCTCCGAGAGGATGAATGAAATGACCCGTACCAAGTCTCTCTTC 1140
Db 1098 ATAAAGCTTACAGCTCCGAGAGGATGAATGAAATGACCCGTACCAAGTCTCTCTTC 1157
QY 1141 TCAGTCAGGTTTAAAGCGTACCCACGAATCCGATGCACTGATCTCTCAAGGCTCA 1200
Db 1158 TCAGTCAGGTTTAAAGCGTACCCACGAATCCGATGCACTGATCTCTCTCAAGGCTCA 1217
QY 1201 TTTCTTTGTGAACAGAGGCTTGGAGGATGGGTACAGCATATCTAAAT 1249
Db 1218 TTTCTTTGTGAACAGAGGCTTGGAGGATGGGTACAGCATATCTCTCAAT 1266

RESULT 3
AI323253
LOCUS
DEFINITION
  mp88b02.y1 Soares thymus 2NbMT Mus musculus cDNA clone IMAGE:576267
  5, similar to gb:U02687 FL CYTOKINE RECEPTOR PRECURSOR (HUMAN);
  gb:X59398 Mouse Flt3 mRNA for tyrosine kinase receptor of the PDGF
  (MOUSE);, mRNA sequence.
ACCESSION
  AI323253
VERSION
  AI323253.1 GI:4057682
KEYWORDS
  EST.
SOURCE
  Mus musculus (house mouse)
  Mus musculus
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
  1 (bases 1 to 765)
REFERENCE
  Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
  Geisler, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
  Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
  Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
  Waterston, R.
  The WashU-HMI Mouse EST Project
  Unpublished
  Contact: Marra M/Mouse EST Project
  WashU-HMI Mouse EST Project
  Washington University School
  4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
  Tel: 314 286 1800
  Fax: 314 286 1810
  Email: mouseest@watson.wustl.edu
  This clone is available royalty-free through LLNL; contact the
  IMAGE Consortium (info@image.llnl.gov) for further information.
  MGI:350915
  This read is a RESEQUENCE of a previously sequenced mouse clone
  This read has been verified (found to hit its original self in the
  correct orientation)
  Seq primer: -40RP from Gibco
  High quality sequence stop: 423.
  Location/Qualifiers
    1..765
      /organism="Mus musculus"
      /mol_type="mRNA"
      /strain="C57BL/6J"
      /db_xref="taxon:10090"
      /clone="IMAGE:576267"
      /sex="male"
      /tissue_type="thymus"
      /dev_stage="4 weeks"
      /lab_host="DH10B"
      /note="Vector: pT73D-Pac (Pharmacia) with a modified
      polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
      was primed with a Not I - oligo(dT) primer [5',
      TGTTCACCATCTGAAGTGGGAGCGCGCTTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGT
      3']; double-stranded cDNA was ligated to Eco RI adaptors
      (Pharmacia), digested with Not I and cloned into the Not I
      and Eco RI sites of the modified pT73 vector. RNA
      provided by Dr. Bertrand Jordan. Library went through two
      rounds of normalization, and was constructed by Bento
      Soares and M.Fatima Bonaldo."
      203 a 176 c 182 g 200 t 4 others
BASE COUNT
ORIGIN
  Query Match 19.0%; Score 655.8; DB 9; Length 765;
  Best Local Similarity 94.6%; Pred. No. 3.2e-74;
  Matches 699; Conservative 0; Mismatches 36; Indels 4; Gaps 2;
  QY 1558 GGGCTTCTCGTCAAAATGCTGTGGTACAAATCTATGGGACGCTTCGGAACCATCTTT 1617
  Db 1 GGGCTTCTCGTCAAAATGCTGTGGTACAAATCTATGGGACGCTTCGGAACCATCTTT 60
  QY 1618 TTAACACTCACCAGGCCCTTCCCTTTTCATCCCAAGACAAATCTCTTATGGACCATT 1677
  Db 61 TTAACACTCACCAGGCCCTTCCCTTTTCATCCCAAGACAAATCTCTTATGGACCATT 120
  QY 1678 GGGCTTCTGCTCCCTTCATTTGTTGTTTCATTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1737
  Db 121 GGGCTTCTGCTCCCTTCATTTGTTGTTTCATTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 180
  QY 1738 CAATTAGTACGAGAGTCACTGTCAGATGATCCAGTGACTGCCCCCTGGATAACGAG 1797
  Db 181 CAATTAGTACGAGAGTCACTGTCAGATGATCCAGTGACTGCCCCCTGGATAACGAG 240
  QY 1798 TACTTCTAGCTTGACTTCAGGGACTATGAATGACTTAAGTGGAGTTCCCGAGAGAG 1857
  Db 241 TACTTCTAGCTTGACTTCAGGGACTATGAATGACTTAAGTGGAGTTCCCGAGAGAG 300
  QY 1858 AACTTAGAGTTTGGGAAGGTCTCTGGGGTCTGGCGCTTTTCGGGAGGGTGTGAACGCCACG 1917
  Db 301 AACTTAGAGTTTGGGAAGGTCTCTGGGGTCTGGCGCTTTTCGGGAGGGTGTGAACGCCACG 360
  QY 1918 GCCTATGGCATTAGTAAACCGGAGTCTCAATTACAGTGGCGGTGAAGATGCTTAAAGAG 1977
```


discovery
 JOURNAL
 MEDLINE
 PUBMED
 COMMENT

Genome Res. 6 (9), 791-806 (1996)
 97044477
 8889548
 Contact: Soares, MB
 Coordinated Laboratory for Computational Genomics
 University of Iowa
 375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 9565
 Email: bento-soares@uiowa.edu
 cDNA Library Preparation: M.B. Soares Lab Clone distribution:
 clones will be available through Research Genetics (www.resgen.com)
 This clone is also available through the I.M.A.G.E. Consortium at
 LNL (info@image.llnl.gov). IMAGE ID= 1793119
 Seq primer: M13 Forward.

FEATURES
 source
 1. .604
 /organism="Rattus norvegicus"
 /mol_type="mRNA"
 /strain="Sprague-Dawley"
 /db_xref="taxon:10116"
 /clone="UI-R-C2p-rg-f-10-0-UI"
 /dev_stage="adult"
 /lab_host="DH10B (Life Technologies)"
 /clone_lib="UI-R-C2p"
 /notes="Vector: pT73D-Pac (Pharmacia) with a modified
 polylinker; Site 1: Not I; Site 2: Eco RI; The UI-R-C2p
 library is a subtracted library derived from the UI-R-C1
 library, which is a subtracted library derived from the
 UI-R-C0 library. The UI-R-C0 library consisted of a
 mixture of individually tagged normalized libraries
 constructed from rat placenta, adult lung, brain, liver,
 kidney, heart, spleen, ovary, muscle, 8, 12 and 18-day
 embryo. The tag is a string of 3-5 nucleotides present
 between the Not I site and the oligo-dT track which allows
 identification of the library of origin of a clone within
 the mixture. The subtracted library (UI-R-C2p) was
 constructed as follows: PCR amplified cDNA inserts from
 UI-R-C1 clones from which 3' ESTs had been derived was
 used as a driver in a hybridization with the UI-R-C1
 library in the form of single-stranded circles. The
 remaining single-stranded circles (subtracted library) was
 purified by hydroxyapatite column chromatography,
 converted to double-stranded circles and electroporated
 into DH10B bacteria (Life Technologies) to generate the
 UI-R-C2p library. This procedure has been previously
 described (Bonaldo, Lennon and Soares, Genome Research 6:
 791-806, 1996)."

BASE COUNT 166 a 132 c 167 g 139 t

ORIGIN

Query Match 15.4%; Score 531.8; DB 10; Length 604;
 Best Local Similarity 93.7%; Pred. No. 2e-58;
 Matches 554; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY 1991 GTGAAAAGAGAGCTCTCATCTGCGAGCTCAAAATGATGACCCACCTGGGACACCATGACA 2050
 DB 8 GTGAAGAGAGGGCGCTCATGGCTGAGCTCAATGATGACCCACCTGGGACCATGACA 67

QY 2051 ACATCGTGAATCTGTGGGGGCGATGCACTGTGAGGGCCAGTGATCTGATTTTGAAT 2110
 DB 68 ACATCGTGAACCTGTGGGGGCGATGCACTGTGAGGGCCAGTGATCTGATTTTGAAT 127

QY 2111 ATGTGTTCTATGTGACCTCTCACTACCTAGAGTAAGAGAGAGATTTTCACAGGA 2170
 DB 128 ATGTGTTGCCATGTGACCTCTCACTACCTAGAGTAAGAGAGAGATTTTCACAGGA 187

QY 2171 CATGGACAGAGATTTTAAAGGAACATAATTTTCACTTCTTACCCTACTTTCCAGGCACATT 2230
 DB 188 CGTGGACAGAGATTTTAAAGGAACATAATTTTCACTTCTTACCCTACTTTCCAGGCACATT 247

QY 2231 CAAATTCAGCATGCTGTTTACGAGAGAGTTCACTGTTTACACCCGCCCTTGGATCAGCTCT 2290

Db 248 CAACTCCAGTATGCCGGTTTCAGAGAGTTTCAGATATACCGCCCTGATCAGGTCT 307
 QY 2291 CAGGGTTCAATGGGAATTCATTCATTCTGAAGATGAGATTGAATATGAAACCAAGAAGA 2350
 Db 308 CAGGGTTCAATGGGAATTCATTCATTCTGAAGATGAGATTGAATATGAAACCAAGAAGA 367
 QY 2351 GGCTGGCAGAGAAGAGAGAGAGAGATTGAAAGCTGTCAGCTTTTGAAGACCTCTTGTCT 2410
 Db 368 GGCTGGGAAGAAGAGAGAGAGAGATTGAAAGCTGTCAGCTTTTGAAGACCTCTTGTCT 427
 QY 2411 TTGGTACCAGTGGCCAAAGGCATGGAATTCCTGGAGTTCAAGTCGTGTGCCACAGAG 2470
 Db 428 TTGGTATCAAGTGGCCAAAGGCATGGAATTCCTGGAGTTCAAGTCGTGTGCCACAGAG 487
 QY 2471 ACCTGGCAGCAGGAATGT 2530
 Db 488 ACCTGGCAGCAGGAATGT 547
 QY 2531 GACTGGCCGAGAGATCTGT 2591
 Db 548 GACTGGCCGAGAGATCTGT 598

RESULT 6
 AI323643/c
 LOCUS
 DEFINITION
 mp8802.x1 Soares thymus 2NbMT Mus musculus cDNA clone IMAGE:576267
 3', similar to gb:U02687 FL CTKINASE RECEPTOR PRECURSOR (HUMAN);
 gb:X59398 Mouse Flt3 mRNA for tyrosine kinase receptor of the PDGF
 (MOUSE);, mRNA sequence.

ACCESSION
 AI323643
 VERSION
 AI323643.1 GI:4058072
 KEYWORDS
 EST.
 SOURCE
 Mus musculus (house mouse)
 ORGANISM
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 747)
 Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
 Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
 Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
 Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
 Waterston, R.
 The WashU-HMI Mouse EST Project
 Unpublished
 Contact: Marra M/Mouse EST Project
 WashU-HMI Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@watson.wustl.edu
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:350915
 This clone was previously sequenced on the 5' end only, this new
 data is from the 3' end
 High quality sequence stop: 432.

FEATURES
 source
 1. .747
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="IMAGE:576267"
 /sex="male"
 /tissue_type="Thymus"
 /dev_stage="4 weeks"
 /lab_host="DH10B"
 /clone_lib="Soares thymus_2NbMT"
 /notes="Vector: pT73D-Pac (Pharmacia) with a modified
 polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA

QY	2637	AATCAAGAGTGACGCTCTGGTCTTACGCACTCCCTCTCTGGAGATATTTTCACTGGGTGT	2696
DB	301	AATCAAGAGTGACGCTCTGGTCTTACGCACTCCCTCTCTGGAGATATTTTCACTGGGTGT	360
QY	2697	GAACCCCTTACCCCTGGCATTCCTCTGCAGCGCTAACTTCTATATAAACTGATTCAGAGTGGATT	2756
DB	361	GAACCCCTTACCCCTGGCATTCCTCTGCAGCGCTAACTTCTATATAAACTGATTCAGAGTGGATT	420
QY	2757	TAAATGGAGCAGCCCATTTCTATGCCAGAGGGATATACCTTTGTATGCAATCCCTGCTG	2816
DB	421	TAAATGGAGCAGCCCATTTCTATGCCAGAGGGATATACCTTTGTATGCAATCCCTGCTG	480
QY	2817	GGCTTTTGACTCAAGGAAG	2835
DB	481	GGCTTTTGACTCAAGGAAG	499

RESULT 8

BB644407

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BB644407 522 bp mRNA linear EST 26-OCT-2001

BB644407 RIKEN full-length enriched, adult male corpora

quadrigemina Mus musculus cDNA clone B230315G04 5', mRNA sequence.

BB644407

BB644407.1 GI:16478944

EST.

Mus musculus

Mus musculus (house mouse)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus;

1 (bases 1 to 522)

Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A.,

Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J., Konno,H., Kouda

M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M.,

Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki

D., Shibata,K., Shingawa,A., Shiraki,T., Sogabe,Y., Suzuki,H.,

Tagamatsu,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toya,T.,

Muramatsu,M. and Hayashizaki,Y.

RIKEN Mouse ESTs (Arakawa,T., et al. 2001)

Unpublished

Contact: Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic

Sciences Center(GSC), Yokohama Institute

The Institute of Physical and Chemical Research (RIKEN)

1-7-22 Suehiro-cho, Teurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Tel: 81-45-503-9222

Fax: 81-45-503-9216

Email: genome-resgsc.riken.go.jp,

URL:http://genome.gsc.riken.go.jp/

Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh

M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.

Normalization and subtraction of cap-trapper-selected cDNAs to

prepare full-length cDNA libraries for rapid discovery of new

genes. Genome Res. 10 (10), 1617-1630 (2000)

wagi,K., Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,

Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura

S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and

Hayashizaki,Y.

RIKEN integrated sequence analysis (RISA) system--384-format

sequencing pipeline with 384 multicapillary sequencer. Genome Res.

10 (11), 1757-1771 (2000)

Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara

Y. and Hayashizaki,Y.

Computer-based methods for the mouse full-length cDNA

encyclopedia: real-time sequence clustering for construction of a

nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)

Kondo,S., Shingawa,A., Saito,T., Kiyosawa,H., Yamanaka,I., Aizawa

K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K. and

Hayashizaki,Y.

Computational Analysis of Full-Length Mouse cDNAs Compared with

Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)

Please visit our web site (<http://genome.gsc.riken.go.jp>) for

further details.

e mouse tissues.

```

RESULT 9
BI360262
LOCUS      BI360262
DEFINITION 387099 MARC 2PIG Sus scrofa cDNA 5', mRNA sequence.
ACCESSION  BI360262
VERSION    BI360262.1
KEYWORDS   EST.
SOURCE     Sus scrofa (pig)
ORGANISM   Sus scrofa
REFERENCE  1 (bases 1 to 621)
AUTHORS   Fahrenkrug, S.C., Smith, T.P.L., Freking, B.A., Cho, J., White, J.,
          Vallst, J., Wise, T., Rohrer, G.A., Pertea, G., Sultana, R., Quackenbush
          , J. and Keele, J.W.
TITLE     Porcine gene discovery by normalized cDNA-library sequencing and
          EST cluster assembly
JOURNAL   Mamm. Genome 13 (8), 475-478 (2002)
MEDLINE   22213789
PUBMED    12226715
COMMENT   Contact: Smith TPL
          USDA, ARS, US Meat Animal Research Center
          PO Box 166, Clay Center, NE 68933-0166, USA
          Tel: 402 762 4366
          Fax: 402 762 4390
          Email: smith@email.marc.usda.gov
          Single pass sequencing. Bases called and alt trimmed with phred
          v0.980904.e. Vector identified by cross_match with the -minscore 18
          and -mismatch 12 options.
          PCR Primers
          FORWARD: AGGAACAGCTATGACCAT
          BACKWARD: GTTTCCAGTCACGACG
          Plate: 139 row: H column: 2
          Seq primer: ATTTAGTGACACTATAG.
FEATURES
    source
        1..621
            /organism="Sus scrofa"
            /mol_type="mRNA"
            /db_xref="taxon:9823"
            /tissue_type="pooled"
            /lab_host="DH10B"
            /clone_lib="MARC 2PIG"
            /notes="Vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI;
            Library made from pooled tissue from testis, ovary,
            endometrium, hypothalamus, pituitary, and placenta."
BASE COUNT  190 a 132 c 141 g 158 t
ORIGIN
Query Match      13.1%; Score 452.6; DB 12; Length 621;
Best Local Similarity 83.2%; Pred. No. 2.4e-48;
Matches 515; Conservative 0; Mismatches 104; Indels 0; Gaps 0;
QY 1702 GTTCTCATTTGTTGATCTCCCAAAATACAAAAGCAATTTAGGTACGAGAGTCAGCTG 1761
      |||||
      3 GTTTAACATGCTAATTTGTCACAGTACAAAAGCGATTCGGTACGAAACCGAGCTG 62
QY 1762 CAGATGATCCAGTGACTGCCGCCCTCGGATACGAGTACTTCTACGTTGACTTCAGGGAC 1821
      |||||
      63 CAGATGGTCCAGTGACGGGGTCCCTGGATACGACTCTTCTACATCGACTTCAGAGAA 122
QY 1822 TATGAATATGACCTTAAGTGGGAGTTCCCGAGAGAGACTTAGAGTTGGGAAGTCTG 1881
      |||||
      123 TATGAATATGACCTCAATGGGAGTTTCCAGGGAAAATTTAGAAATTTGGGAAGTTCTG 182
QY 1882 GGTCTCGGCGCTTTCCGGGGGATGACGCCACCGCTATGCGATTAGTAAACGGGA 1941
      |||||
      183 GGATCCGGCGCTTTTGGAAAGATGATGACGCAACTGCCTATGGAATCAGTAAACTGGA 242
QY 1942 GTCTCAATTCCAGTGGCGGTGAAGATGCTTAAAGAGAGAAAGCTGACAGCTGTGAAAAAGAA 2001
      |||||
      243 GTGTCAATCCAGTGGCAGTCGCAAAATGCTGAAGAAAGAGCAGACAGCTCGGAACGAG 302
QY 2002 GCTCTCATGTCGGAGCTCAAAATGATGACCCACCTGGGACACCATGACAAACATCGTGAAT 2061

```

```

Db 303 GCTCTCATGTTGATCTCAAAATGATGACCACCTGGCGAGCCAGAGATATAGTGAAC 362
QY 2062 CTGCTGGGGCAGTCACACACTGTGAGGGCCAGTGTACTTGTGATTTTGAATATTTGTTCTAT 2121
      |||||
      363 CTGCTGGGGCCTGCACCCCTGTGAGGGCCCATTTACTTGTGATTTTGAATATTTGTTCTAC 422
QY 2122 GGTGACCTCTCACTACCTAAAGAGTAAAGAGAGAGAGTTTTCACAGACATGACAGAG 2181
      |||||
      423 GGTGATCTTCTCACTATCTAAAGAGTAAAGAGAGAGAAATTTTCATAGACATGACCGAG 482
QY 2182 ATTTTAAAGAACATATTTTCACTTCTTACCTTCTTCCAGGCACATTCCTCAATTCAGC 2241
      |||||
      483 ATTTTCAAGGAACATATTTTCACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 542
QY 2242 ATGCTCTGTTTCCAGAGAGTTTCACTACACCGCCCTTGGATCAGCTCTCAGGGTTCAT 2301
      |||||
      543 ATGCGGGTTCAAGAGAGTTTCAATACACCGACATCGGATCTCTCTCAGGATTCAT 602
QY 2302 GGGAAATTCATTCATCTG 2320
      |||||
      603 GGGAAATTCATTCATCTG 621
      |||||

RESULT 10
BY727686
LOCUS      BY727686
DEFINITION BY727686 RIKEN full-length enriched, 6 days neonate medulla
          oblongata Mus musculus cDNA clone B73004D22 5', mRNA sequence.
ACCESSION  BY727686
VERSION    BY727686.1
KEYWORDS   EST.
SOURCE     Mus musculus (house mouse)
ORGANISM   Mus musculus
REFERENCE  1 (bases 1 to 658)
AUTHORS   Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,
          Nikaido, I., Osato, N., Saito, R., Suzuki, H., Yamana, I., Kiyosawa, H.,
          Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C.,
          Gojobori, T., Baldarelli, R., Hill, D.P., Bult, C., Hume, D.A.,
          Quackenbush, J., Schriml, L.M., Kanapin, A., Mateuda, H., Batalov, S.,
          Beisel, K.W., Blake, J.A., Bradt, D., Brusic, V., Chothia, C., Corbani
          , L.E., Cougins, S., Dalla, E., Dragani, T.A., Fletcher, C.F., Forrest
          , A., Frazer, K.S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A.,
          Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, I.J.,
          Jarvis, E.D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedzierski, R.M.,
          King, B.L., Konagaya, A., Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons
          , P.A., Maglott, D.R., Maltais, L., Marchionni, L., McKenzie, L., Miki
          , H., Nagashima, T., Numata, K., Okido, T., Pavan, W.J., Pertea, G.,
          Pesole, G., Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D.,
          Ramachandran, S., Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring
          , B.Z., Ringwald, M., Sandelin, A., Schneider, C., Semple, C.A., Setou
          , M., Shmida, K., Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale
          , R.D., Tomita, M., Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y.,
          Watanabe, Y., Wells, C., Wilming, L.G., Wynshaw-Boris, A., Yanagisawa
          , M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A.,
          Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura
          , M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Nakamura
          , A., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii
          , Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata
          , K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander
          , E.S., Rogers, J., Birney, E. and Havaehizaki, Y.
          Analysis of the mouse transcriptome based on functional annotation
          of 60,770 full-length cDNAs
          Nature 420, 563-573 (2002)
JOURNAL    22354683
MEDLINE    12466851
PUBMED     12466851
COMMENT    Contact: Yoshihide Havaehizaki
          Laboratory for Genome Exploration Research Group, RIKEN Genomic
          Sciences Center (GSC), Yokohama Institute
          The Institute of Physical and Chemical Research (RIKEN)
          1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

```

Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.go.jp,
URL: <http://genome.gsc.riken.go.jp/>
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hashizume, W., Hayashida, K., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Koya, S., Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R., Ohno, K., Ohsato, N., Saito, R., Sakazume, M., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Takeda, Y., Waki, K., Watahiki, A., Muramatsu, M. and Hayashizaki, Y.: Direct Submission

Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)

Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)

RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences' Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site (<http://genome.gsc.riken.go.jp/>) for further details.

```

Location/Qualifiers
1..658
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="B730044D22"
/tissue_type="medulla oblongata"
/dev_stage="6 days neonate"
/lab_host="DH10B"
/clone_lib="RIKEN full-length enriched, 6 days neonate
medulla oblongata"
/note="Site 1: SalI; Site 2: BamHI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN. Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5'
GAGAGAGAGAGGATCCAGAGACTCTTTTITTTTTTTTNN 3'], cDNA was
prepared by using trihalose thermo-activated reverse
transcriptase and subsequently enriched for full-length by
cap-trapper. cDNA went through one round of normalization
to Rot = 20.0 and subtraction to Rot = 459.0. Second
strand cDNA was prepared with the primer adapter of
sequence [5' GAGAGAGAGATTCGAGGTAATTAATTAATCCCCCCCCCCCC
3']. cDNA was cleaved with XhoI and BamHI. Vector: a
modified pBluescript KS(+) after bulk excision from Lambda
phiC11"

```

	126 a	202 c	204 g	124 t	2 others
h. Similarity	12.9%	Score 446.6;	DB 14;	Length 658;	
S2; Conservative	97.6%;	Pred. No. 1.3e-47;			
	0;	Mismatches 11;	Indels 0;	Gaps 0;	
1	CGGCGCTGGCTACCGCGCGCTCCGAGAGGCCATGCGGGCGTTGGCGGAGCGACGCGACGG	60			
6	CGGCGCTGGCTACCGCGCGCTCCGAGAGGCCATGCGGGCGTTGGCGGAGCGACGCGACGG	255			
1	CGGCGTGTGCTGTGTTGTTTGTTCAGTAAATGATCTTGAGACCGTTCAAAACCAAGAC	120			
	CGGCGTGTGCTGTGTTGTTTGTTCAGTAAATGATCTTGAGACCGTTCAAAACCAAGAC	315			

QY	121	CTGCCGTGATCAAGTGTCTTTTAATCACTCATGAGAA	CAATGGCTCATCAGCGGAAAG	180
Db	316	CTGCTGTGATCAAGTGTCTTTTAATCACTCATGAGAA	CAATGGCTCATCAGCGGAAAG	375
QY	181	CCATCATCTTACCGAATGTCGGAGGATCCCAGAAGAC	CTCCAGTGTACCCCGAGCGC	240
Db	376	CCATCATCTGTCGGAATGTCGGAGGATCCCAGAAGAC	CTCCAGTGTGCCCCGAGCGC	435
QY	241	CAGAGTGAAGGACGGTATATGAAGCGGCACCGTGGAG	GTGGCCGAGTGTGGTCCATC	300
Db	436	CAGAGTGAAGGACGGTATATGAAGCGGCACCGTGGAG	GTGGCCGAGTGTGGTCCATC	495
QY	301	ACCTGCAAGTCAGCTGCCACCCAGGGGACCTTTCTG	CTGCTCTGGGTCTTTAAGCAC	360
Db	496	ACCTGCAAGTCAGCTGCCACCCAGGGGACCTTTCTG	CTGCTCTGGGTCTTTAAGCAC	555
QY	361	AGCTCCCTGGGTGCCAGCGGCATTTTGATTTTACAAA	CAGAGGAATCGTTTCCATGGCC	420
Db	556	AGCTCCCTGGGTGTCAGCGGCATTTTGATTTTACAAA	CAGAGGAATCGTTTCCATGGCC	615
QY	421	ATCTTGAACGTACAGAGACCCAGGAGGAGAATACCT	ACTCC	463
Db	616	CATCTTGAGTGCACAGAGACCCAGGAGGAGAATAC	CTACTCC	658

RESULT 11
BE651447/c
LOCUS
DEFINITION

BE651447
UI-M-BH3-ack-d-06-0-UI-r1 NIH BMAP M.S4 Mus musculus cDNA clone
UT-M-BH3-ack-d-06-0-UI 5' mRNA sequence

445 bp mRNA linear EST 06-SEP-2000

UI-M-BH3-atk-d-06-0-UI 5', mRNA sequence.
 BE651447
 BE651447.1 GI:9977271
 EST.
 Mus musculus (house mouse)
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 445)
 Bonaldo,M.F., Lennon,G. and Soares,M.B.
 Normalization and subtraction: two approaches to facilitate gene
 discovery
 Genome Res. 6 (9), 791-806 (1996)
 97044477
 8889548
 Contact: Chin, H
 National Institute of Mental Health
 6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
 20892-9643, USA
 Tel: 301 443 1706
 Fax: 301 443 9890
 Email: MEST@mail.nih.gov
 CDNA Library Preparation: M.B. Soares Lab Clone distribution:
 Researchers may obtain BMAP cDNA clones from RESEARCH GENETICS. It
 should be noted that Bento Soares is generating a small number of
 additional specialized non-redundant arrays of BMAP cDNAs whose
 availability will be considered under appropriate and limited
 collaborative arrangements
 Seq primer: M13 Reverse.
 Location/Qualifiers
 1..445
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UI-M-BH3-atk-d-06-0-UI"
 /dev_stages="27-32 days"
 /lab_host="DH10B (Life Technologies)"
 /clone_lib="NIH BMAP M S4"
 /note="Vector: p773D-Pac (Pharmacia) with a modified
 polylinker; Site_1: Not I; Site_2: Eco RI; The
 NIH BMAP M S4 library is a subtracted library of a series,
 ultimately derived from a mixture of individually tagged
 normalized libraries from ten regions of the mouse brain

(cerebellum, brain stems, olfactory bulbs, hypothalamus, cortex, amygdala, basal ganglia, pineal gland, striatum, hippocampus) after a series of subtractions to reduce the representation of cDNAs from which ESTs had already been generated. The following serially subtracted libraries were generated in this process: NIH_BMAP_M_S4, NIH_BMAP_M_S3.3, NIH_BMAP_M_S3.2, NIH_BMAP_M_S3.1, NIH_BMAP_M_S2, NIH_BMAP_M_S1. The subtracted library (NIH_BMAP_M_S4) was constructed as follows: PCR-amplified cDNA inserts from NIH_BMAP_M_S3.3, NIH_BMAP_M_S3.2, and NIH_BMAP_M_S3.1 clones from which 3' ESTs had been derived was used as a driver in a hybridization with a pool of the NIH_BMAP_M_S3.3, NIH_BMAP_M_S3.2, and NIH_BMAP_M_S3.1 libraries in the form of single-stranded circles. The remaining single-stranded circles (subtracted library) was purified by hydroxyapatite column chromatography, converted to double-stranded circles and electroporated into DH10B bacteria (LifeTechnologies) to generate the NIH_BMAP_M_S4 library. This procedure has been previously described (Bonaldo, Lennon and Soares, Genome Research 6:791-806, 1996)".

BASE COUNT	115 t	123 c
96 a	111 g	123 c

Query Match.	12.9%;	Score 445;	DB 10;	Length 445;
Best Local Similarity	100.0%;	Prod. No. 2.6e-47;		
Matches 445;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	612	TGTGGAGTGGGTGCTCTGCGAGCTCCCA	CAGGGAAGCTGTAA	CAAGAGGCCCTGCTGT 671
Db	445	TGTGGAGTGGGTGCTCTGCGAGCTCCCA	CAGGGAAGCTGTAA	CAAGAGGCCCTGCTGT 386
Qy	672	TGTCAGAAAGGAGGAAAGGTACTTTCAT	GTAGTTCTTCGGAACAGACAT	TCAGATGCTGTGC 731
Db	385	TGT CAGAAAGGAGGAAAGGTACTTTCAT	GTAGTTCTTCGGAACAGACAT	TCAGATGCTGTGC 326
Qy	732	TAGAAATGCACTGGGCGCGGAATCCACA	CAAGCTGTTCA	CACATAGATCTAAACAGGCTCC 791
Db	325	TAGAAATGCACTGGGCGCGGAATGCA	CAAGCTGTTCA	CACATAGATCTAAACAGGCTCC 266
Qy	792	TCAGAGCACACTGCCGCCAGTTATTCCT	GTAAAGTGGGGAACCC	TTGTGGATCAGGTGTA 851
Db	265	TCAGAGCACACTGCCGCCAGTTATTCCT	GTAAAGTGGGGAACCC	TTGTGGATCAGGTGTA 206
Qy	852	GGCCATCCATGTGAACCATGATTTCGGG	CTCACCCTGGGAGCTGGAACAGCAAGGCC	CTGGGA 911
Db	205	GGCCATCCATGTGAACCATGATTTCGGG	CTCACCCTGGGAGCTGGAACAGCAAGGCC	CTGGGA 146
Qy	912	GGAGGGCAGCTACTTTGAGATGAGTAC	TCTCCACAAACAGGACCATGAT	TCGGATTCT 971
Db	145	GGAGGGCAGCTACTTTGAGATGAGTAC	TCTCCACAAACAGGACCATGAT	TCGGATTCT 86
Qy	972	CTTTGGCCCTTTGTGTTCCGTGGGAAG	GAACAGCACCCGGATATTACACCTGCT	CTTCCCTC 1031
Db	85	CTTTGGCCCTTTGTGTTCCGTGGGAAG	GAACAGCACCCGGATATTACACCTGCT	CTTCCCTC 26
Qy	1032	AAAGCACCCCGCAGTCAGCGTTG	1056	
Db	25	AAAGCACCCCGCAGTCAGCGTTG	1	

RESULT 12	BI461248	714 bp	linear	EST 21-AUG-2001
LOCUS	BI461248	603206574F1	NIH_MGC_97	Homo sapiens cDNA clone IMAGE:5272266 5',
DEFINITION	mRNA sequence.			
ACCESSION	BI461248			
VERSION	BI461248.1	GI:15251904		
KEYWORDS	EST.			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.			

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

1 (pages 1 to 714)

NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shizuo
Toshiyuki and Piero Carninci (RIKEN)
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM11687 row: c column: 19
High quality sequence spot: 709.

FEATURES	SOURCE
1. Geographical Location: The study area is located in the northern part of the state, covering an area of approximately 100 square miles.	State Department of Transportation
2. Population: The population of the study area is estimated to be around 50,000 people, with a significant portion being young adults.	U.S. Census Bureau
3. Economic Activity: The primary economic activity in the area is manufacturing, with several large industrial plants operating.	Local Chamber of Commerce
4. Infrastructure: The area is well-served by major highways and has a growing network of public transportation options.	Metropolitan Planning Organization
5. Environmental Concerns: There are ongoing concerns about air quality and water resources, particularly in the industrial zones.	Environmental Protection Agency
6. Social Issues: The area faces challenges related to income inequality and access to healthcare services.	Local Health Department
7. Education: The area has several public schools and a few private institutions, with a focus on STEM education.	Local School District
8. Healthcare: There is a growing need for healthcare facilities, particularly in the area of mental health services.	Local Health Department
9. Transportation: The area is experiencing increased traffic congestion, particularly during peak hours.	State Department of Transportation
10. Community Development: There are ongoing efforts to improve the quality of life and attract new businesses to the area.	Local Chamber of Commerce

```

1. .714
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:527266"
/lab_host="DH10B"
/clone_lib="NTH MGC 97"

```

/clone_id= NIH_MGC_3",
 "note=Organ: testis; Vector: pBluescriptR (modified
 pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI (Gtcgag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTWN-3',
 size-selected for average insert size 2.2 kb and
 normalized to ROT 5. This is a primary library enriched
 for full-length clones and constructed using the
 Cap-trapper method (Carninci, in preparation). Library
 constructed by M. Brownstein (NIMH/NHGRI, National
 Institutes of Health). Note: this is a NIH MGC Library."

BASE COUNT	182 a	171 c	191 q	170 t

[illegible]

Db 535 ACCAATTACACATATTTGTTACAGTGGATTAAGAAATACCTGCTTTACACATTAAGA 594
 Qy 541 AGACCTTACTTTAGGAGATGGAACACAGATGCACTGCTGCACTCTCGAGGCTTT 600
 Db 595 AGACCTTACTTTAGAAAAATGGAACACAGAGCGCTGCTGCTGCACTATCTGAGAGCTT 654
 Qy 601 CCGAGCCCACTGAGAGTGGGCTCTGAGCTCCACAGGGAAGCTGTAAGAGAGAA 660
 Db 655 CCAGAGCCGATCGTGAATGGGCTTTTGCATTCACAGGGGGAAGCTGTAAGAGAGCA 714

RESULT 13
 BU695104 597 bp mRNA linear EST 08-OCT-2002
 LOCUS LL2in11905T7 Hematopoietic Stem Cell Subtracted Library Mus
 DEFINITION musculus cDNA 5' similar to flk-2, mRNA sequence.
 ACCESSION BU695104
 VERSION BU695104.1 GI:23596587
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 597)
 Phillips, R.L.; Ernst, R.E.; Brunk, B.P.; Ivanova, N.; Mahan, M.A.;
 Deanehan, J.K.; Moore, K.A.; Overton, G.C. and Lemischka, I.R.
 The genetic program of hematopoietic stem cells
 Science 288 (5471), 1635-1640 (2000)
 20295303
 MEDLINE 10834841
 PUBMED
 CONTACT: Lemischka, Ihor R.
 Department of Molecular Biology
 Princeton University
 Lewis Thomas Laboratory, Washington Road, Princeton, NJ 08544, USA
 Tel: 609 258 2838
 Fax: 609 258 2759
 Email: ilemischka@molbio.princeton.edu
 These ESTs are derived from a subtracted cDNA library enriched for
 gene products expressed in day 14-14.5 fetal liver hematopoietic
 stem cells defined as Lineageneg/lo, AA4.1pos, ckitpos, ly6A/E
 (Sca-1)pos
 Seq primer: M13Reverse or T7.
 Location/Qualifiers
 1. 597
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57Bl/6J"
 /db_xref="taxon:10090"
 /tissue_type="Hematopoietic"
 /cell_type="Stem Cells; Lineageneg/lo, AA4.1pos, ckitpos,
 ly6A/E (Sca-1)pos"
 /dev_stage="Embryonic day 14-14.5"
 /lab_host="DH10B"
 /clone_lib="Hematopoietic Stem Cell Subtracted Library"
 /note="Organ: Fetal Liver; Vector: Sport 1; Site 1: Sal I;
 Site 2: Not I; Two directionally cloned cDNA libraries
 were made from fetal liver hematopoietic stem cells
 enriched to be Lineageneg/lo, AA4.1pos, ckitpos, ly6A/E
 (Sca-1)pos called Scapos and from AA4.1neg fetal liver
 cells. Subtractive hybridization was performed by
 hybridization of the target, Scapos, single stranded cDNA
 library in pSport1 to biotinylated RNA transcribed from
 the driver, AA4.1neg cDNA library in pSport2 with inserts
 cloned in the complementary orientation. For detailed
 protocols and additional information please see our
 website at http://stemcell.princeton.edu."

BASE COUNT 136 a 140 c 152 g 131 t 38 others
 ORIGIN
 Query Match 12.4%; Score 428.4; DB 13; Length 597;
 Best Local Similarity 92.3%; Pred. No. 2.9e-45;
 Matches 469; Conservative 0; Mismatches 35; Indels 4; Gaps 3;

Qy 724 TGCTGTCTAGAAATGCACTGGGCGGAAATGCAACCAAGCTGTTCACCATAGATCTAAAC 783
 Db 2 TGCTGTCTAGAAATGCACTGGGCGGAAATGCAACCAAGCTGTTCACCATAGATCTAAAC 61
 Qy 784 CAGCTCTCTAGAGCACACTGCCCGAGTATCTCTGAAAGTGGGGGACCCCTTGTGGATC 843
 Db 62 CAGCTCTCTAGAGCACACTGCCCGAGTATCTCTGAAAGTGGGGGACCCCTTGTGGATC 121
 Qy 844 AGTGTAAAGCCATCCATGTGAACCATGGATTTCGGGCTCACTGGGAGCTGGAACACAA 903
 Db 122 AGTGTAAAGCCATCCATGTGAACCATGGATTTCGGGCTCACTGGGAGCTGGAACACAA 181
 Qy 904 GCCCTGAGGAGGCGAGCTACTTTGAGATGATGATCTACTCCACAAACAGGACCATGATT 963
 Db 182 GCCCTGAGGAGGCGAGCTACTTTGAGATGATGATCTACTCCACAAACAGGACCATGATT 241
 Qy 964 CGGATCTCTTGGCTTTGTCTTCCTGGGAAGAACGACACCGGATATACACCTGC 1023
 Db 242 CGGATCTCTTGGCTTTGTCTTCCTGGGAAGAACGACACCGGATATACACCTGC 301
 Qy 1024 TCTTCTCAAGACACCCAGCCAGTGTGGTGTGACCATCTCTAGAAAGGGTTTATA 1083
 Db 302 TCTTCTCAAGACACCCAGCCAGTGTGGTGTGACCATCTCTAGAAAGGGTTTATA 361
 Qy 1084 AACCTACCACTCGCAAGAGAGTATGAAATGACCCGTACGAAAGTCTCTCTCTCA 1143
 Db 362 AACCTATCAGTCCGAGAGAGTATGAAATGACCCGTACGAAAGTCTCTCTCTCA 421
 Qy 1144 GTCAGGTTT-AAAGCGTACCCAGCAATCCGATGACGTTGG--ATCTTCTCAAGCTCA 1200
 Db 422 GTCAGGTTTAAAGCGTACCCAGCAATCCGATGACGTTGGGATCTTCTCTNAAGCTCA 481
 Qy 1201 -TTTCTTGTGAACAGAGCGCTGGAG 1227
 Db 482 TTTTCTTGTGANNNGANAGGCGCTGGG 509

RESULT 14
 BQ556177 515 bp mRNA linear EST 20-JUN-2002
 LOCUS H4038808-5 NIA Mouse 7.4K cDNA Clone Set Mus musculus cDNA clone
 DEFINITION H4038808 5', mRNA sequence.
 ACCESSION BQ556177
 VERSION BQ556177.1 GI:21457065
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 515)
 VanBuren, V.; Piao, Y.; Dudekula, D.B.; Qian, Y.; Carter, M.G.; Martin,
 P.R.; Stagg, C.A.; Bassey, U.; Aiba, K.; Hamatani, T.; Kargul, G.J.,
 Luo, A.G.; Kelsio, J.; Hide, W. and Ko, M.S.H.
 Assembly, verification, and initial annotation of NIA 7.4K mouse
 cDNA clone set
 Genome Res. 12 (12), 1999-2003 (2002)
 MEDLINE 22354164
 PUBMED 12466305
 COMMENT Other ESTs: H4038808-3
 Contact: Yong Qian
 Laboratory of Genetics
 National Institute on Aging/National Institutes of Health
 333 Cassell Drive, Suite 3000, Baltimore, MD 21224-6820, USA
 Email: cdna@igsun.grc.nia.nih.gov
 This clone set has been freely distributed to the community. Please
 visit http://igsun.grc.nia.nih.gov/cDNA/NIA_7.4k.html for details.
 Plate: H4038 row: E column: 08
 Seq primer: -21M13 Reverse
 High quality sequence stop: 515
 POLYA=No.
 Location/Qualifiers
 1. 515

FEATURES
 source

```
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="niaEST:H4038E08-5"
/clone="H4038E08"
/sex="mixed"
/dev_stage="mixed"
/lab_host="DH10B"
/clone_lib="NIA Mouse 7.4K cDNA Clone Set"
/notes="Vector: pSPORT1; Site 1: SalI; Site 2: NotI; This clone is among a rearrayed set of 7,407 clones from more than 20 cDNA libraries."
BASE COUNT      139 a 138 c 128 g 110 t
ORIGIN

Query Match      12.3%; Score 426.2; DB 13; Length 515;
Best Local Similarity 92.8%; Pred. No. 5.9e-45;
Matches 478; Conservative 0; Mismatches 13; Indels 24; Gaps 2;

QY 2725 GCTAACTTCTATAAAGTATTGAGAGTGGATTTAAATGAGAGAGCCATTCATGACCA 2784
      |||||
Db 1 GCTAACTTCTATAAAGTATTGAGAGTGGATTTAAATGAGAGAGCCATTCATGACCA 60
      |||||

QY 2785 GAAGGGATATCTTCTTAATGCAATCTGCTGGGCTTTTGACTCAAGGAAGCGGCATCC 2844
      |||||
Db 61 GAAGGGATATCTTCTTAATGCAATCTGCTGGGCTTTTGACTCAAGGAAGCGGCATCC 120
      |||||

QY 2845 TTCCCCAAGCTGACTTCATTTTGGAGTGTGAGCTGGCAGAGGAGCAAGAGC----- 2897
      |||||
Db 121 TTCCCCAAGCTGACTTCATTTTGGAGTGTGAGCTGGCAGAGGAGCAAGAGCGATGAT 180
      |||||

QY 2898 -----ATGATCAGAACATCTCATCTCATCTACCAAAACAGCGGCGCCCTC 2941
      |||||
Db 181 CAGAACATGGGTGGCAACGCTCCAGAACATCCATCTATACCAAAACAGCGGCGCCCTC 240
      |||||

QY 2942 AGCAGAGAGCGGGCTCAGAGCC-CAGTCGCCACAGCGCCAGGTGAAGATTCAACAGAA 3000
      |||||
Db 241 AGCAGAGAGCGGGCTCAGAGCCGCCATCGCCACAGCCAGGTGAAGATTCAACAGAA 300
      |||||

QY 3001 AGAAGTTAGCGAGAGCGCTTGACCCCGCCAGCTAGCAGGCTGTAGACCGCAGAGCCA 3060
      |||||
Db 301 AGAAGTTAGCGAGAGCGCTTGACCCCGCCAGCTAGCAGGCTGTAGACCGCAGAGCCA 360
      |||||

QY 3061 AGATTAGCTCGCTCTGAGGAAGCGCCTAGCGCGGTTCGCTGGAGCTTTCTCT 3120
      |||||
Db 361 AGATTAGCTCGCTCTGAGGAAGCGCCTACAGCGCGGTTCGCTGGAGCTTTCTCT 420
      |||||

QY 3121 AGATGCTGTCTGCATTTACTCCAAAGTGACTTCTATAAAATCAAACTCTCTCGCACAG 3180
      |||||
Db 421 AGATGCTGTCTGCATTTACTCCAAAGTGACTTCTATAAAATCAAACTCTCTCGCACAG 480
      |||||

QY 3181 GCGGGAGAGCCCAATATGAGACTTGTGTGAGCC 3215
      |||||
Db 481 GTGGGAGAGCCCAATATGAGACTTGTGTGAGCC 515
      |||||

RESULT 15
LOCUS      BM117950/c
DEFINITION 403 bp mRNA linear EST 09-JUN-2003
            L0858E03-3 NIA Mouse Newborn Brain cDNA Library Mus musculus cDNA
            clone L0858E03 3', mRNA sequence.
ACCESSION  BM117950
VERSION     1
KEYWORDS   EST.
SOURCE     Mus musculus (house mouse)
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
            1 (bases 1 to 403)
REFERENCE  Piao, Y., Ko, N.T., Lim, M.K. and Ko, M.S.H.
            Construction of long-transcript enriched cDNA libraries from
            submicrogram amounts of total RNAs by a universal PCR amplification
```

method
Genome Res. 11 (9), 1553-1558 (2001)
21429098
11544199
On Nov 26, 2001 this sequence version replaced gi:17080968.
Contact: Dawood B. Dudekula
Laboratory of Genetics
National Institute on Aging/National Institutes of Health
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
Email: cdna@leuun.grc.nia.nih.gov
Plate: L0858 row: E column: 03
Seq primer: -21M13 Forward
High quality sequence stop: 403
POLYA=Yes.

Location/Qualifiers
1..403
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="niaEST:L0858E03-3"
/db_xref="taxon:10090"
/clone="L0858E03"
/tissue_type="Newborn Brain"
/dev_stage="Newborn"
/lab_host="DH10B"
/clone_lib="NIA Mouse Newborn Brain cDNA Library"
/note="Organ: Brain; Vector: pSPORT1 (Invitrogen); Site 1:
SalI; Site 2: NotI; Mouse cDNA project by the Laboratory
of Genetics, National Institute on Aging (NIA), Intramural
Research Program, NIH (http://lgsun.grc.nia.nih.gov/cDNA).
Double-stranded cDNAs were synthesized with an Oligo(dT)
primer [Invitrogen: 5'-
pACATGTTCTAGATCGGAGCGCGCCCTTTTCTTTTCTTTT-3'] from 48
microgram of total RNA, treated with T4 DNA polymerase,
and purified by ethanol-precipitation. The cDNAs were
ligated to Lone-linker Li-Sal3 (Ref. Development
127:1737-1749 (2000) [PMID:10725249]), purified by
phenol/chloroform, and separated from free linkers by
Centricon 100. Then, the cDNAs were digested with SalI and
NotI enzymes, and cloned into SalI/NotI site of pSPORT1
plasmid vector. The DH10B E. Coli host was transformed
with ligation mixture by the standard chemical method. The
average insert size is about 1.9 kb. The library was
constructed by Yulan Piao (NIA)."

BASE COUNT 118 a 83 c 89 g 113 t

Query Match 10.5%; Score 361.6; DB 12; Length 403;
Best Local Similarity 95.5%; Pred. No. 1.1e-36;
Matches 384; Conservative 0; Mismatches 10; Indels 8; Gaps 1;

QY 3034 CTTAGCAGGCTGTAGACCCAGAGCCAGATTAGCTCGCCTCTGAGGAAGCGCCCTACA 3093
 |||||
Db 403 CTTAGCAGGCTGTAGACCCAGAGCCAGATTAGCTCGCCTCTGAGGAAGCGCCCTACA 344
 |||||

QY 3094 GCGGCTGTGCTCGTGGAGCTTTCTCTAGATGCTGTCTGCCATTACTCCAAAGTGATTC 3153
 |||||
Db 343 GCGGCTGTGCTCGTGGAGCTTTCTCTAGATGCTGTCTGCCATTACTCCAAAGTGATTC 284
 |||||

QY 3154 TATAAATCAAACTCTCTCTCGCACAGGCGGAGAGCCAAATATGAGACTTGTGTGTGAG 3213
 |||||
Db 283 TATAAATCAAACTCTCTCTCGCACAGGTTGGAGAGCCAAATATGAGACTTGTGTGTGAG 224
 |||||

QY 3214 CCGGCTACCTTGGGGGCGCTT-----TCCACGAGCTTGGGGGAAAGCCATGTATCT 3265
 |||||
Db 223 CCGGCTACCTTGGGGGCGCTT-----TCCAGGCGCCCGCCAGGCTTGGGGAAAGCCATGTATCT 164
 |||||

QY 3266 GAAATATAGTATATCTTGTAAATAGCTGAAACAAACCCGTTTGTGTAAGGGA 3325
 |||||
Db 163 GAAATATAGTATATCTTGTAAATAGCTGAAACAAACCCGTTTGTGTAAGGGA 104
 |||||

QY 3326 AAGCTTAAATATGATTTTAAATATCTATGTTTAAATATCTATGTAACCTTTTTCATCTAT 3385
 |||||

Db 103 AAGCTAAATATGATTTTAAATAATCTATGTTTAAATACTATGTAACTTTTTCATCTAT 44

Qy 3386 TTAGTGATATATTTTATGGATGGAATAAACTTTCCTACTGTA 3427

Db 43 TTAGTGATATATTTTATGGATGGAATAAACTTTCCTACTGTA 2

Search completed: August 28, 2003, 05:42:11
Job time : 8349.07 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

QOM nucleic - nucleic search, using sw model

Run on: August 27, 2003, 14:22:21 ; Search time 1028.85 Seconds
(without alignments)
9059.793 Million cell updates

Title: US-09-919-408A-1
Perfect score: 3453
Sequence: 1 GCGGCTGGCTACCGCGGC.....AAAAAAAAAAAAAAAAAAAA 3453

Scoring table: IDENTITY_NUC
Gapop 10.0 . Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_19Jun03:*

Database :		N_GeneSeq_19Jun03:
1:	/SIDS1/cgdata/geneseq/geneSeq-emb1/NA1980.DAT:	
2:	/SIDS1/cgdata/geneseq/geneSeq-emb1/NA1981.DAT:	*
3:	/SIDS1/cgdata/geneseq/geneSeq-emb1/NA1982.DAT:	*
4:	/SIDS1/cgdata/geneseq/geneSeq-emb1/NA1983.DAT:	*
5:	/SIDS1/cgdata/geneseq/geneSeq-emb1/NA1984.DAT:	*
6:	/SIDS1/cgdata/geneseq/geneSeq-emb1/NA1985.DAT:	*
7:	/SIDS1/cgdata/geneseq/geneSeq-emb1/NA1986.DAT:	*
8:	/SIDS1/cgdata/geneseq/geneSeq-emb1/NA1987.DAT:	*
9:	/SIDS1/cgdata/geneseq/geneSeq-emb1/NA1988.DAT:	*
10:	/SIDS1/cgdata/geneseq/geneSeq-emb1/NA1989.DAT:	*
11:	/SIDS1/cgdata/geneseq/geneSeq-emb1/NA1990.DAT:	*
12:	/SIDS1/cgdata/geneseq/geneSeq-emb1/NA1991.DAT:	*
13:	/SIDS1/cgdata/geneseq/geneSeq-emb1/NA1992.DAT:	*
14:	/SIDS1/cgdata/geneseq/geneSeq-emb1/NA1993.DAT:	*
15:	/SIDS1/cgdata/geneseq/geneSeq-emb1/NA1994.DAT:	*
16:	/SIDS1/cgdata/geneseq/geneSeq-emb1/NA1995.DAT:	*
17:	/SIDS1/cgdata/geneseq/geneSeq-emb1/NA1996.DAT:	*
18:	/SIDS1/cgdata/geneseq/geneSeq-emb1/NA1997.DAT:	*
19:	/SIDS1/cgdata/geneseq/geneSeq-emb1/NA1998.DAT:	*
20:	/SIDS1/cgdata/geneseq/geneSeq-emb1/NA1999.DAT:	*
21:	/SIDS1/cgdata/geneseq/geneSeq-emb1/NA2000.DAT:	*
22:	/SIDS1/cgdata/geneseq/geneSeq-emb1/NA2001A.DAT:	*
23:	/SIDS1/cgdata/geneseq/geneSeq-emb1/NA2001B.DAT:	*
24:	/SIDS1/cgdata/geneseq/geneSeq-emb1/NA2002.DAT:	*
25:	/SIDS1/cgdata/geneseq/geneSeq-emb1/NA2003.DAT:	*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query #	Score	Match	Length	DB ID	Description
1	3453	100.0	3453	16	AAQ81012	Flk2 receptor prot
2	3453	100.0	3453	17	AAT38733	Human foetal liver
3	3453	100.0	3453	18	AAT72118	Murine flk-2 recep
4	3453	100.0	3453	20	ABY77514	Murine flk-2 cDNA.
5	3453	100.0	3453	24	ABX55044	Mouse cDNA encodin
6	3451.4	100.0	3453	24	AAQ53502	Murine flk-2 cDNA.
7	3451.4	100.0	3453	14	AAQ35249	Murine flk-2 codin
8	3451.4	100.0	3453	15	AAQ79068	Mouse flk-2 cDNA.

9	3448.2	99.9	3453	13	AAQ29954	Murine flk-2 cDNA
10	3446.6	99.8	3453	14	AAQ40914	Murine flk-2 cDNA
11	3357	97.2	3440	24	AAQ24483	Murine receptor pr
12	3342.4	96.8	3521	16	AAQ00801	Flk2/flt3 tyrosine
13	2227.2	64.5	3501	16	AAQ81013	Flk2 receptor prot
14	2227.2	64.5	3501	16	AAQ79069	Human flk-2 cDNA
15	2227.2	64.5	3501	17	AAQ38734	Murine foetal live
16	2227.2	64.5	3501	18	AAQ72117	Human flk-2 recept
17	2227.2	64.5	3501	20	AAQ77515	Human flk-2 cDNA
18	2227.2	64.5	3501	24	ABX55045	Human cDNA encodin
19	2227.2	64.5	3501	24	ABX55045	Human flk-2 cDNA
20	2225.6	64.5	3501	14	AAQ53503	Human flk-2 cDNA
21	2225.6	64.5	3501	14	AAQ40915	Human flk-2 cDNA
22	2225.6	64.5	3501	14	AAQ35250	Human flk-2 coding
23	2225.4	64.4	3501	14	AAQ35250	Human receptor pro
24	2220.4	63.9	3489	24	AAQ42484	Human STR-1 cDNA
25	2204.8	63.9	3476	16	AAQ91536	Human Flk2/flt3 ty
26	2198.2	63.7	3475	16	AAQ00802	pTK gene LpTK25
27	2085	60.4	3120	14	AAQ49756	Protein tyrosine-k
28	2081.8	60.3	3120	16	AAQ03096	Human receptor typ
29	2024.6	58.6	2949	19	AAV39041	Human receptor typ
30	2018.8	58.5	2958	19	AAV39042	Human receptor typ
31	2017.2	58.4	2958	19	AAV39040	Human receptor typ
32	1994.8	57.8	2982	19	AAV39039	Human receptor typ
33	1984.2	57.5	2978	19	AAV39038	Human receptor typ
34	1199.6	34.7	1894	15	AAQ54036	Flk-2w5 gene Mus
35	1119.6	32.4	2247	19	AAV52294	Sequence pMON32390
36	357.4	10.4	3992	23	AAV79666	DNA encoding novel
37	357.4	10.4	3992	23	AAV79666	Breast carcinoma r
38	357.4	9.6	3069	21	AZ444718	Human macrophage c
39	357.4	10.4	3992	24	ABK48105	Human gene express
40	357.4	10.4	3992	24	ABK48105	Bovine c-Kit bK-1
41	357.4	10.4	3992	24	ABK48105	Hampshire porcine
42	337.4	9.6	3069	21	AZ444718	Human c-kit oncoge
43	331.4	9.6	2919	20	AAV80687	Human Kit/stem cel
44	329.9	9.5	5084	19	AAV20443	Breast cancer rela
45	329.9	9.5	5084	22	ABL13426	Ovary cancer relat
46	329.9	9.5	5084	24	ABL64113	DNA encoding novel
47	329.9	9.5	5084	24	ABL68085	DNA encoding novel
48	325.4	9.4	6390	23	AAV79665	DNA encoding novel
49	325.4	9.4	6390	23	AAV79665	DNA encoding novel
50	325.4	9.4	6390	23	AAV79665	DNA encoding novel

ALIGNMENTS

	RESULT 1	
XX	AAQ81012	
ID	AAQ81012 standard; cDNA; 3453 BP.	
XX	AC	
XX	AAQ81012;	
DT	(updated)	
DD	25-MAR-2003	
TT	16-AUG-1995 (first entry)	
DE		
XE	Flk2 receptor protein-tyrosine-kinase cDNA.	
KX		
KW	Mouse Flk2; receptor protein-tyrosine-kinase; primitive hematopoietic cell; fetal liver kinase; ds.	
OS	Mus musculus.	
FF		
FH	Key	Location/Qualifiers
CDS	31..3009	/tag= a
FT		/product= flk2 receptor protein-tyrosine-kinase
FT	sigs_peptide	31..111
FT		/tag= b
FT	mat_peptide	112..3006
FT		/tag= c
PN	WO9500554-A2.	
XP		
PD	05-JAN-1995.	
PX		
PF	17-JUN-1994;	94WO-USO6944.
PP		
XX		

Db	1681	CTCTGTCTCCCTTCATATGTGTCTCTCATTTGTGTTGATCTGCCACAAATATACAAAAGCA	1740
Qy	1741	TTTAGGTACGAGAGTCAGCTGCAGATGATCCAGSTGACTGGCCCCCTGGATAACGAGTAC	1800
Db	1741	TTTTAGGTACGAGAGTCAGCTGCAGATGATCCAGTGTACTGGCCCCCTGGATAACGAGTAC	1800
Qy	1801	TTCTACGTTGACTTCAGGGACTATGATATATGACCTTAAAGTGGGAGTTCCCGAGAGAGAC	1860
Db	1801	TTCTACGTTGACTTCAGGGACTATGATATATGACCTTAAAGTGGGAGTTCCCGAGAGAGAC	1860
Qy	1861	TTAGAGTTTGGGAAGGTCCTGGGGTCTGGCGCTTCGGGAGGGTGATGACGCCACGGCC	1920
Db	1861	TTAGAGTTTGGGAAGGTCCTGGGGTCTGGCGCTTCGGGAGGGTGATGACGCCACGGCC	1920
Qy	1921	TATGGCATTAGTAAAAACGGGAGTCTCAATTCAGGTGGCGGTGAAGATGCTTAAAGAGAAA	1980
Db	1921	TATGGCATTAGTAAAAACGGGAGTCTCAATTCAGGTGGCGGTGAAGATGCTTAAAGAGAAA	1980
Qy	1981	GCTGACAGCTGTGAAAAAAGAGTCTCATGTCCGAGCTCAAAATGATGACCCACCTGGGA	2040
Db	1981	GCTGACAGCTGTGAAAAAAGAGTCTCATGTCCGAGCTCAAAATGATGACCCACCTGGGA	2040
Qy	2041	CACCATGACACATCGTGAATCTGCTGGGGCATGCACATGTTCAGGGCCAGTGTACTTTCG	2100
Db	2041	CACCATGACACATCGTGAATCTGCTGGGGCATGCACATGTTCAGGGCCAGTGTACTTTCG	2100
Qy	2101	ATTTTTCAATATTTGTTGCTATGTGTGACCTCCTCAACTACCTAAAGAAAGTAAAGAGAGAAG	2160
Db	2101	ATTTTTGAAATATTTGTTGCTATGTGTGACCTCCTCAACTACCTAAAGAAAGTAAAGAGAGAAG	2160
Qy	2161	TTTTCACAGGACATGGAACAGAGATTTTAAAGAACATAATTTTCAGTTTACCCTACTTTTC	2220
Db	2161	TTTTCACAGGACATGGAACAGAGATTTTAAAGAACATAATTTTCAGTTTACCCTACTTTTC	2220
Qy	2221	CAGGCACATTCAAATTCAGCATGCTGTTTCACGAGAAGTTCAGTTTACACCCGCCCTTG	2280
Db	2221	CAGGCACATTCAAATTCAGCATGCTGTTTCACGAGAAGTTCAGTTTACACCCGCCCTTG	2280
Qy	2281	GATCAGCTCTCAGGGTTCAATGGGAATTCAAATTCATTCGAAATGAGATGAGATGAATGAA	2340
Db	2281	GATCAGCTCTCAGGGTTCAATGGGAATTCAAATTCATTCGAAATGAGATGAGATGAATGAA	2340
Qy	2341	AACCAGAAGAGGCTGGCAGAGAAGAGAGAGAGAGATTTGAACGTCTGACGTTTGAAGAC	2400
Db	2341	AACCAGAAGAGGCTGGCAGAGAAGAGAGAGAGAGATTTGAACGTCTGACGTTTGAAGAC	2400
Qy	2401	CTCCTTTGCTTTGGTACCAAGTGGCCAAAGGCATGGAAATTCCTCGAGTTTCAAGTCGTGT	2460
Db	2401	CTCCTTTGCTTTGGTACCAAGTGGCCAAAGGCATGGAAATTCCTCGAGTTTCAAGTCGTGT	2460
Qy	2461	GTCCACAGAGACCTGGCAGCCAGGAATGTGTGTGTCCACCCACGGGAAGGTGGTGAAGATC	2520
Db	2461	GTCCACAGAGACCTGGCAGCCAGGAATGTGTGTGTCCACCCACGGGAAGGTGGTGAAGATC	2520
Qy	2521	TGTGACATTTTGGACTGGCCCCGAGACATCTTGACGCACTCCAGCTACGTCTCAGGGGCAAC	2580
Db	2521	TGTGACATTTTGGACTGGCCCCGAGACATCTTGACGCACTCCAGCTACGTCTCAGGGGCAAC	2580
Qy	2581	GCACGGCTGCCGGTGAAGTGATGGCACCCGAGAGCTTATTTGAAGGGATCTACACATTC	2640
Db	2581	GCACGGCTGCCGGTGAAGTGATGGCACCCGAGAGCTTATTTGAAGGGATCTACACATTC	2640
Qy	2641	AAGAGTGACGCTCTGCTCTACGGCATCTTCTCTGGGAGATATTTTCACTGGGTGTGAAC	2700
Db	2641	AAGAGTGACGCTCTGCTCTACGGCATCTTCTCTGGGAGATATTTTCACTGGGTGTGAAC	2700
Qy	2701	CCTTACCTTGGCATTTCTGTGACGCTAACTTCTATATAAATGATTCAGAGTGGATTTTAAA	2760
Db	2701	CCTTACCTTGGCATTTCTGTGACGCTAACTTCTATATAAATGATTCAGAGTGGATTTTAAA	2760
Qy	2761	ATGGAGAGGCCATTCTATGCCACAGAAAGGATATACTTTGTAAATGCATCTCTGTGGCT	2820
Db	2761	ATGGAGAGGCCATTCTATGCCACAGAAAGGATATACTTTGTAAATGCATCTCTGTGGCT	2820

Qy	2821	TTTGACTCAAGGAAGCGGCCATCTTCTCCCAACCTGACTTCATTTTACGATGTCAGCTG	2880
Db	2821	TTTGGACTCAAGGAAGCGGCCATCTTCTCCCAACCTGACTTCATTTTACGATGTCAGCTG	2880
Qy	2881	GCAGAGCGAGAGCAAGCATGTATCAGAAACATCCATCCATCTACCAAAAACAGGCGGCCCT	2940
Db	2881	GCAGAGCGAGAGAGCATGTATCAGAAACATCCATCCATCTACCAAAAACAGGCGGCCCT	2940
Qy	2941	CAGCAGAGAGCGGGCTCAGAGCCCAAGTCGCCACAGCGCCAGGTGAAGATTACAGAGAA	3000
Db	2941	CAGCAGAGAGCGGGCTCAGAGCCCAAGTCGCCACAGCGCCAGGTGAAGATTACAGAGAA	3000
Qy	3001	AGAACTTAGCGAGAGGCGCTTGACCCCGCCACCTAGCAGGCTGTAGACCGCAGAGCCA	3060
Db	3001	AGAAGTTAGCGAGAGGCGCTTGACCCCGCCACCTAGCAGGCTGTAGACCGCAGAGCCA	3060
Qy	3061	AGATTAGCTCGGCTCTTGAGGAAGCGCCCTACAGCGCTTGCTTCGCTGGACTTTTCTCT	3120
Db	3061	AGATTAGCTCGGCTCTTGAGGAAGCGCCCTACAGCGCTTGCTTCGCTGGACTTTTCTCT	3120
Qy	3121	AGATGCTGTCGCCATTAATCTCAAAGTGACTTCTATAAATCAAACTCTCCTCGCACAG	3180
Db	3121	AGATGCTGTCGCCATTAATCTCAAAGTGACTTCTATAAATCAAACTCTCCTCGCACAG	3180
Qy	3181	CGCGGAGAGCCAATAATAGACACTTGTGTGTGAGCGCCCTACCTGGGGGCCCTTTCACAG	3240
Db	3181	CGCGGAGAGCCAATAATAGACACTTGTGTGTGAGCGCCCTACCTGGGGGCCCTTTCACAG	3240
Qy	3241	AGCTTGAGGGGAAAGCCATGTATCTGAAATATAGTATATTTCTGTGTAATAACGTGGAACAA	3300
Db	3241	AGCTTGAGGGGAAAGCCATGTATCTGAAATATAGTATATTTCTGTGTAATAACGTGGAACAA	3300
Qy	3301	ACCAAAACCGTTTTTTTCTTAAGGAAGCTTAATATGATTTTTTAAAAATCTATGTTTTAA	3360
Db	3301	ACCAAAACCGTTTTTTCTTAAGGAAGCTTAATATGATTTTTTAAAAATCTATGTTTTAA	3360
Qy	3361	AATACTATGTAACTTTTTTTCATCTATTATTAGTGATATATTTTTATGGATGGAAATAAAGTTTC	3420
Db	3361	AATACTATGTAACTTTTTTTCATCTATTATTAGTGATATATTTTTATGGATGGAAATAAAGTTTC	3420
Qy	3421	TACTGTAAAAAAAAAAAAAAAAAAAAAAAAAA 3453	
Db	3421	TACTGTAAAAAAAAAAAAAAAAAAAAAAAAAA 3453	

RESULT 2
AAT38733
ID AAT38733 standard; cDNA; 3453 BP.

XX	AAT38733;
AC	
XX	25-MAR-2003 (updated)
DT	11-DEC-1996 (first entry)
DT	
XX	Human foetal liver kinase 2 cDNA.
DE	
XX	Human; foetal liver kinase 2; flk-2; protein tyrosine kinase;
KW	monoclonal; antibody; extracellular domain; receptor assay;
KW	haematopoietic stem cell; ligand; stimulation; proliferation;
KW	differentiation; treatment; anaemia; bone marrow damage;
KW	cancer chemotherapy; radiation; ds.

XX	OS	Homo sapiens.	Location/Qualifiers
XX	FX	Key	31..3009
FX	FT	CDS	/*tag= a
FT	FT		31..111
FT	FT	sig_peptide	/*tag= b
FT	FT		112..3006
FT	FT	mat_peptide	/*tag= c
FX	FX		

Qy	1261	AAGAACAGCCAGGAGAGTACATATTTCTATGCAGAAAAATGATGACGCCCAAGTTCACAAA	1320
Db	1261	AAGAACAGCCAGGAGAGTACATATTTCTATGCAGAAAAATGATGACGCCCAAGTTCACAAA	1320
Qy	1321	ATGTTCAAGCGTGAATATAAGAAAGAAACCTCAAGTGTCTAGCAAAATGCTCAGCCAGCCAG	1380
Db	1321	ATGTTCAAGCGTGAATATAAGAAAGAAACCTCAAGTGTCTAGCAAAATGCTCAGCCAGCCAG	1380
Qy	1381	CGCTCTGTTTCTCTGATGGCTACCCGCTACCCCTCTTTGGACCTTGGAAAGAAAGTGTTCGGAC	1440
Db	1381	CGCTCTGTTTCTCTGATGGCTACCCGCTACCCCTCTTTGGACCTTGGAAAGAAAGTGTTCGGAC	1440
Qy	1441	AAATCTCCCAATTGCACGGAGGAAATCCACAGAGGAGTTTGGAAATAAAAAGCTTAACAGA	1500
Db	1441	AAATCTCCCAATTGCACGGAGGAAATCCACAGAGGAGTTTGGAAATAAAAAGCTTAACAGA	1500
Qy	1501	AAAGTGTTCGGCAGTGGGTGTCACAGACAGTACTCTAAATATGAGTGAAGCCGGGAAAGGG	1560
Db	1501	AAAGTGTTCGGCAGTGGGTGTCACAGACAGTACTCTAAATATGAGTGAAGCCGGGAAAGGG	1560
Qy	1561	CTTCTGCTCAAAATGCTGTGCGTCAAAATTTATGCGCACGCTCTTCGCAAAACCATCTTTTAA	1620
Db	1561	CTTCTGCTCAAAATGCTGTGCGTCAAAATTTATGCGCACGCTCTTCGCAAAACCATCTTTTAA	1620
Qy	1621	AACTCACAGGCCCTTCCCTTTTCATCAAGACAAATCTCTTCTATATGGAACCATTTGGG	1680
Db	1621	AACTCACAGGCCCTTCCCTTTTCATCAAGACAAATCTCTTCTATATGGAACCATTTGGG	1680
Qy	1681	CTCTGTCTCCCTTCATATGTTGTTCTCATTTGTGTTGATCTGCCACAATACAAAAGCAA	1740
Db	1681	CTCTGTCTCCCTTCATATGTTGTTCTCATTTGTGTTGATCTGCCACAATACAAAAGCAA	1740
Qy	1741	TTTATGTTACGAGAGTCAAGTGCAGATGATCCAGGTGACTGCGCCCTCTGGATTAACGAGTAC	1800
Db	1741	TTTATGTTACGAGAGTCAAGTGCAGATGATCCAGGTGACTGCGCCCTCTGGATTAACGAGTAC	1800
Qy	1801	TTCTACGTTTGAATATAAAGGAGTCTCAATTCAGGTGGCGGTGATGAACGAGAGAAAC	1860
Db	1801	TTCTACGTTTGAATATAAAGGAGTCTCAATTCAGGTGGCGGTGATGAACGAGAGAAAC	1860
Qy	1861	TTAGAGTTTGGGAAGGTCCTGGGTCCTGGCGCTTTCCGGAGGGTGATGAACGAGAGAAAC	1920
Db	1861	TTAGAGTTTGGGAAGGTCCTGGGTCCTGGCGCTTTCCGGAGGGTGATGAACGAGAGAAAC	1920
Qy	1921	TATGGCATTAGTAAAAAGGAGTCTCAATTCAGGTGGCGGTGAAAGATGCTTAAAGAGAAA	1980
Db	1921	TATGGCATTAGTAAAAAGGAGTCTCAATTCAGGTGGCGGTGAAAGATGCTTAAAGAGAAA	1980
Qy	1981	GCTGACAGCTGTGAAAAAGAGCTCTCATGTCGGAGTCTCAAAATGATGACCCACCTGGGA	2040
Db	1981	GCTGACAGCTGTGAAAAAGAGCTCTCATGTCGGAGTCTCAAAATGATGACCCACCTGGGA	2040
Qy	2041	CACCATGACAACTCGTGAATCTCTCTGGGGCATGCACTGTTCAGGGCCAGTGTACTTG	2100
Db	2041	CACCATGACAACTCGTGAATCTCTCTGGGGCATGCACTGTTCAGGGCCAGTGTACTTG	2100
Qy	2101	ATTTTTGAAATATGTTGCTATGTTGACCTCTCAACTACCTTAAGAAAGTAAAGAGAGAAG	2160
Db	2101	ATTTTTGAAATATGTTGCTATGTTGACCTCTCAACTACCTTAAGAAAGTAAAGAGAGAAG	2160
Qy	2161	TTTTCACAGGACATGACAGAGATTTTAAAGGAAACATAATTTTCACTTACCTTACTTTC	2220
Db	2161	TTTTCACAGGACATGACAGAGATTTTAAAGGAAACATAATTTTCACTTACCTTACTTTC	2220
Qy	2221	CAGGCAATTTCAAATTCACAGATGCTGGTTTCACAGAGTTTCAGTTTACCCCTTACTTTC	2280
Db	2221	CAGGCAATTTCAAATTCACAGATGCTGGTTTCACAGAGTTTCAGTTTACCCCTTACTTTC	2280
Qy	2281	GATCAGCTCTCAGAGTTTCAATGGGAATTCATTTCAATCTGAAGATGAGATTGAATATGAA	2340
Db	2281	GATCAGCTCTCAGAGTTTCAATGGGAATTCATTTCAATCTGAAGATGAGATTGAATATGAA	2340
Qy	2341	AACCAGAAAGAGGCTGGCAGAGAAAGAGGAGAGATTTTGAAACGTGTCAGCTTTTGAAGAC	2400

Db	2341	AA	CCAGAAAGAGGCTGGCAGAAAGAGAGGAGAAATTTGAACTGCTGACGTTTGAAGAC	2400
Qy	2401	CT	CCCTTTGCTTTTCGTACCAAGTGGCCAAAGGCATGGAATTCCTCGAGTTTCAAGTCGTGT	2460
Db	2401	CT	CCCTTTGCTTTTCGTACCAAGTGGCCAAAGGCATGGAATTCCTCGAGTTTCAAGTCGTGT	2460
Qy	2461	GT	CCACAGAGACCTGGCAGCAGGAATGTGTTGGTCAACCCACGGGAAGTGGTGAAGATC	2520
Db	2461	GT	CCACAGAGACCTGGCAGCAGGAATGTGTTGGTCAACCCACGGGAAGTGGTGAAGATC	2520
Qy	2521	TG	TGACCTTTGGACTGGCCCGCAGACATCCTCAGCGGACTCCAGCTACGTCGTCAGGGGCAAC	2580
Db	2521	TG	TGACCTTTGGACTGGCCCGCAGACATCCTCAGCGGACTCCAGCTACGTCGTCAGGGGCAAC	2580
Qy	2581	GC	CGGCTGCGGCTGAAGTGGATGGCACCCGAGAGCTTATTTGAAGGATCTTACACAATC	2640
Db	2581	GC	CGGCTGCGGCTGAAGTGGATGGCACCCGAGAGCTTATTTGAAGGATCTTACACAATC	2640
Qy	2641	AA	GAGTGAAGTCTGGTCTTACGGCATCCTTCTCTGGGAGATATTTTCACTGGGTGTGAAC	2700
Db	2641	AA	GAGTGAAGTCTGGTCTTACGGCATCCTTCTCTGGGAGATATTTTCACTGGGTGTGAAC	2700
Qy	2701	CC	TACCTGGCATTTCTGTGCGCGTAACTTCTATAAACTGATTCAGAGTGGATTTAAA	2760
Db	2701	CC	TACCTGGCATTTCTGTGCGCGTAACTTCTATAAACTGATTCAGAGTGGATTTAAA	2760
Qy	2761	AT	GAGCAGAGCATTTCTATGCGCACAGAAGGATATACCTTCTGATGCAATCCTGCTGGCT	2820
Db	2761	AT	GAGCAGAGCATTTCTATGCGCACAGAAGGATATACCTTCTGATGCAATCCTGCTGGCT	2820
Qy	2821	TT	TGACTCAAGGAAGCGGCCATCCTTCCCAAACCTGACTTCAITTTTGGATGTGAGCTG	2880
Db	2821	TT	TGACTCAAGGAAGCGGCCATCCTTCCCAAACCTGACTTCAITTTTGGATGTGAGCTG	2880
Qy	2881	GC	AGGACAGAGAGCATGTATCAGAACTCCATCCATCTACCAAACAGGGGGCCCT	2940
Db	2881	GC	AGGACAGAGAGCATGTATCAGAACTCCATCCATCTACCAAACAGGGGGCCCT	2940
Qy	2941	CAG	CAGAGAGCGGGCTCAGAGCCAGTCGCGCACAGGCCAGGTGAAGATTACAGAGAA	3000
Db	2941	CAG	CAGAGAGCGGGCTCAGAGCCAGTCGCGCACAGGCCAGGTGAAGATTACAGAGAA	3000
Qy	3001	AGA	AGTTAGCGAGAGCGCTTGGACCCCGCACCCCTAGCAGGCTGTAGACCGCAGAGCCA	3060
Db	3001	AGA	AGTTAGCGAGAGCGCTTGGACCCCGCACCCCTAGCAGGCTGTAGACCGCAGAGCCA	3060
Qy	3061	AG	ATTAGCTGCTCTGAGGAGCGCCCTACAGCGGTTGCTTCGCTGGACATTTCCTCT	3120
Db	3061	AG	ATTAGCTGCTCTGAGGAGCGCCCTACAGCGGTTGCTTCGCTGGACATTTCCTCT	3120
Qy	3121	AG	ATGCTGTCTGCCATTACTTCAAGTGAATTTATATAAACTCAAACTCTCTCGCACAG	3180
Db	3121	AG	ATGCTGTCTGCCATTACTTCAAGTGAATTTATATAAACTCAAACTCTCTCGCACAG	3180
Qy	3181	GC	GGAGAGCCAAATATAGACTTTGTTGGTGGAGCCGCTACCCCTGGGGCCCTTTCACG	3240
Db	3181	GC	GGAGAGCCAAATATAGACTTTGTTGGTGGAGCCGCTACCCCTGGGGCCCTTTCACG	3240
Qy	3241	AG	CTTGAGGGGAAGCCATGTATCTGAATAATAGTATATTTCTTGTAATACTGTAACAA	3300
Db	3241	AG	CTTGAGGGGAAGCCATGTATCTGAATAATAGTATATTTCTTGTAATACTGTAACAA	3300
Qy	3301	AC	CAAAACCCGTTTTTCTGAAGGAAAGCTAAATATGATTTTTTAAAACTATGTTTTAA	3360
Db	3301	AC	CAAAACCCGTTTTTCTGAAGGAAAGCTAAATATGATTTTTTAAAACTATGTTTTAA	3360
Qy	3361	AA	TACTATGTAACTTTTTTCATCTATTTAGTGATATATTTTATGATGGAATAAATTTTC	3420
Db	3361	AA	TACTATGTAACTTTTTTCATCTATTTAGTGATATATTTTATGATGGAATAAATTTTC	3420
Qy	3421	TAC	TGTAAAAAATAAAAAAAAAAAAAAAAAAAAAA	3453

Db 3421 TACTGTAAAAA... 3453

RESULT 4
 AAX77514
 ID AAX77514 standard; cDNA; 3453 BP.
 XX
 AC AAX77514;
 XX
 XX 05-AUG-1999 (first entry)
 XX
 DE Murine flk-2 cDNA.
 XX
 XX Murine; flk-2; flk-1; cell isolation; fetal liver kinase; receptor;
 KW monoclonal; polyclonal; antibody; tyrosine kinase; ds.
 XX
 XX Mus sp.
 XX
 XX Key Location/Qualifiers
 FH CDS 31..3009
 FT FT /*tag= a
 FT FT /product= "flk-2"
 XX
 PN US5912133-A.
 XX
 XX 15-JUN-1999.
 XX
 XX 10-FEB-1998; 98US-0021324.
 XX
 XX 19-NOV-1992; 92US-0977451.
 PR 02-APR-1991; 91US-0679666.
 PR 28-JUN-1991; 91US-0728913.
 PR 15-NOV-1991; 91US-0793065.
 PR 24-DEC-1991; 91US-0813593.
 PR 26-JUN-1992; 92US-0906397.
 PR 12-NOV-1992; 92US-0975049.
 PR 30-APR-1993; 93US-0055269.
 PR 31-OCT-1994; 94US-0252498.
 PR 15-FEB-1996; 96US-0601891.
 XX
 PA (UYPR-) UNIV PRINCETON.
 XX
 XX Lemischka IR;
 XX
 XX WPI: 1999-357194/30.
 DR P-PSDB; AAY08616.
 XX
 PT Isolating hematopoietic cells expressing fetal liver kinase 1
 PT receptors
 XX
 XX Disclosure; Fig 1a; 59pp; English.
 XX
 CC This invention describes a novel method of isolating cells expressing
 CC fetal liver kinase 1 (flk-1) receptors on their surface and comprises
 CC binding the cells to a polyclonal or monoclonal antibody specific to
 CC the FLK-1 receptor and isolating the cells that have bound to the
 CC antibody. The method can be used to isolate hematopoietic stem cells in
 CC any mammal but preferably a rat, mouse, rabbit or human. The proteins of
 CC the invention belong to the receptor protein family. This sequence
 CC encodes the murine flk-2 protein which is used in the method of the
 CC invention.
 XX
 SQ Sequence 3453 BP; 947 A; 822 C; 875 G; 809 T; 0 other;

Query Match 100.0%; Score 3453; DB 20; Length 3453;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 3453; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGGCTGGCTACCGGGCTCCGAGGCCATGCGGGCTTGGCGAGCGACGACCGG 60
 Db 1 CGGGCTGGCTACCGGGCTCCGAGGCCATGCGGGCTTGGCGAGCGACGACCGG 60
 Qy 61. CGGCTGCTGCTGTTGTTTGTTCAGTAATGTTTGTGAGACCGTTACAAACCAAGAC 120

Db 61 CGGCTGCTGCTGCTGTTGTTTGTTCAGTAATGTTTGTGAGACCGTTACAAACCAAGAC 120
 Qy 121 CTGCTGTGATCAAGTGTGTTTAAATCAGTCATGAGAACAAATGCTCATCGCGGGAAG 180
 Db 121 CTGCTGTGATCAAGTGTGTTTAAATCAGTCATGAGAACAAATGCTCATCGCGGGAAG 180
 Qy 181 CCATCATGTAACCAATGTTGCGAGGATCCCGAGAGACCTCCAGTGTACCCGAGGCGC 240
 Db 181 CCATCATGTAACCAATGTTGCGAGGATCCCGAGAGACCTCCAGTGTACCCGAGGCGC 240
 Qy 241 CAGAGTGAAGGACGGTATATGAAGCGGCAACCGTGGAGGTGGCGGATCTGGGTCCATC 300
 Db 241 CAGAGTGAAGGACGGTATATGAAGCGGCAACCGTGGAGGTGGCGGATCTGGGTCCATC 300
 Qy 301 ACCCTGCAAGTGCAGCTCGCCACCCAGGGGACCTTTCTCCCTCTGGGTCTTTAAGCAC 360
 Db 301 ACCCTGCAAGTGCAGCTCGCCACCCAGGGGACCTTTCTCCCTCTGGGTCTTTAAGCAC 360
 Qy 361 AGTCTCCTGGGCTGCCAGCCGCACTTTGATTTACAAACAGAGGAATCGTTTCCATGGCC 420
 Db 361 AGTCTCCTGGGCTGCCAGCCGCACTTTGATTTACAAACAGAGGAATCGTTTCCATGGCC 420
 Qy 421 ATCTTGAACGTGACAGACCCAGGCAAGGAGATACCTACTCCATATTCAGAGCGAACGC 480
 Db 421 ATCTTGAACGTGACAGACCCAGGCAAGGAGATACCTACTCCATATTCAGAGCGAACGC 480
 Qy 481 GCCAATCTACACAGTACTGTTTACAGTGAATGTAAGAGATACACAGCTGTATGCTTAAGG 540
 Db 481 GCCAATCTACACAGTACTGTTTACAGTGAATGTAAGAGATACACAGCTGTATGCTTAAGG 540
 Qy 541 AGACCTTACTTTAGGAAGATGGAACCCAGGATGCACTGCTCTGCATCTCCGAGGGTGT 600
 Db 541 AGACCTTACTTTAGGAAGATGGAACCCAGGATGCACTGCTCTGCATCTCCGAGGGTGT 600
 Qy 601 CCGAGCCCACTGTGAGTGGGTGCTCTGAGCTCCACAGGAAAGCTGTAAAGAGAA 660
 Db 601 CCGAGCCCACTGTGAGTGGGTGCTCTGAGCTCCACAGGAAAGCTGTAAAGAGAA 660
 Qy 661 GGCCCTGCTGTTGTGAGAAAGGAGGAAGGATGTTCTATGAGTGTTCGGAACAGACATC 720
 Db 661 GGCCCTGCTGTTGTGAGAAAGGAGGAAGGATGTTCTATGAGTGTTCGGAACAGACATC 720
 Qy 721 AGATGCTGTCTAGAAATGCACTGGGCGCGCAATGCACCAAGCTGTTTACCATAGATCTA 780
 Db 721 AGATGCTGTCTAGAAATGCACTGGGCGCGCAATGCACCAAGCTGTTTACCATAGATCTA 780
 Qy 781 AACCAGGCTCTCAGAGCACACTGCCCCAGTTTCTGAAAGTGGGGGAAACCTTTGTGG 840
 Db 781 AACCAGGCTCTCAGAGCACACTGCCCCAGTTTCTGAAAGTGGGGGAAACCTTTGTGG 840
 Qy 841 ATCAGGTGTAAGGCCATCCATGTGACCATGATTCGGGCTACCTGGGAGCTGGGAAGAC 900
 Db 841 ATCAGGTGTAAGGCCATCCATGTGACCATGATTCGGGCTACCTGGGAGCTGGGAAGAC 900
 Qy 901 AAAGCCCTGGAGGAGGAGCTACTTTGAGATGAGTACTTCTCCACAAACAGACCATG 960
 Db 901 AAAGCCCTGGAGGAGGAGCTACTTTGAGATGAGTACTTCTCCACAAACAGACCATG 960
 Qy 961 ATTCCGATTTCTTTGGGCTTTGTTCTTCCCGTGGGAAGGAACGACACCGGATATTACAC 1020
 Db 961 ATTCCGATTTCTTTGGGCTTTGTTCTTCCCGTGGGAAGGAACGACACCGGATATTACAC 1020
 Qy 1021 TGCTCTTCTTCAAGCACCCAGCCAGCTGAGGTTGGTGACCATCTCTAGAAAAAGGGTTT 1080
 Db 1021 TGCTCTTCTTCAAGCACCCAGCCAGCTGAGGTTGGTGACCATCTCTAGAAAAAGGGTTT 1080
 Qy 1081 ATAAACGCTACCGCTCGCAAGAGATGATGAAATTTGACCCCTACGAAAGTCTTCTGCTTC 1140
 Db 1081 ATAAACGCTACCGCTCGCAAGAGATGATGAAATTTGACCCCTACGAAAGTCTTCTGCTTC 1140
 Qy 1141 TCAGTCAAGTTTAAAGCGGTACCCAGAAATCCGATCCGATCGAGTGGATCTTCTCTCAAGCCTCA 1200

Qy	721	AGATGCTGTGCTAGAAATGCACTGGGCGCGAAATGCACCAAGCTGTTTACCATAGATCTTA	780
Db	721	AGATGCTGTGCTAGAAATGCACTGGGCGCGAATGCACCAAGCTGTTTACCATAGATCTTA	780
Qy	781	AACCGGCTCCTCAGAGCACACTGCCCCAGTTTATTCCTGAAAGTGGGGAAACCTTTGGG	840
Db	781	AACCGGCTCCTCAGAGCACACTGCCCCAGTTTATTCCTGAAAGTGGGGAAACCTTTGGG	840
Qy	841	ATCAGTGTGAAGCCATCATGTGAACCATGTGATTCGGGCTCACCTGGAGCTGGAAGAC	900
Db	841	ATCAGTGTGAAGCCATCCATGTGAACCATGTGATTCGGGCTCACCTGGAGCTGGAAGAC	900
Qy	901	AAAGCCCTGGAGGAGGCGAGCTACTTTTGAGATGAGTACTCTCCACAAAACAGGACCATG	960
Db	901	AAAGCCCTGGAGGAGGCGAGCTACTTTTGAGATGAGTACTCTCCACAAAACAGGACCATG	960
Qy	961	ATTTCGGATTCTCTTTGGCCCTTTGTCTCCGTGGGAAGGAACGACACCGGATATTACACC	1020
Db	961	ATTTCGGATTCTCTTTGGCCCTTTGTCTCCGTGGGAAGGAACGACACCGGATATTACACC	1020
Qy	1021	TGCTCTTCTCTCAAGACACCCAGCCAGTCAGCGTTGGTGACCATCTTAGAAAAAGGGTTT	1080
Db	1021	TGCTCTTCTCTCAAGACACCCAGCCAGTCAGCGTTGGTGACCATCTTAGAAAAAGGGTTT	1080
Qy	1081	ATAAAGCTACCAAGCTCGCAAGAAGATGAAATTTGACCCGTAACGAAAGTTCTGCTTC	1140
Db	1081	ATAAAGCTACCAAGCTCGCAAGAAGATGAAATTTGACCCGTAACGAAAGTTCTGCTTC	1140
Qy	1141	TCAGTCAGGTTTAAAGGTTACCAACGATCCGATGCACTGGGATCTTCTCTCAAGCCCTCA	1200
Db	1141	TCAGTCAGGTTTAAAGGTTACCAACGATCCGATGCACTGGGATCTTCTCTCAAGCCCTCA	1200
Qy	1201	TTTTCTTTGTAACAGAGAGGCTTGGAGATGGGTACAGCATATCTAAATTTTTCGATCAT	1260
Db	1201	TTTTCTTTGTAACAGAGAGGCTTGGAGATGGGTACAGCATATCTAAATTTTTCGATCAT	1260
Qy	1261	AAGAACAGCCAGAGAGATACATATCTATGCAAGAAATGATGACCCGATTTCCACCAA	1320
Db	1261	AAGAACAGCCAGAGAGATACATATCTATGCAAGAAATGATGACCCGATTTCCACCAA	1320
Qy	1321	ATGTTACGCTGAATATAAAGAAAGAACCTCAAGTGTGTAGCAATGGCTCAGCAGCCAG	1380
Db	1321	ATGTTACGCTGAATATAAAGAAAGAACCTCAAGTGTGTAGCAATGGCTCAGCAGCCAG	1380
Qy	1381	CGGTCTGTTCTCTGTATGGCTACCCGCTACCCCTTTGGACCTTGGAGAAAGTGTTCGGAC	1440
Db	1381	CGGTCTGTTCTCTGTATGGCTACCCGCTACCCCTTTGGACCTTGGAGAAAGTGTTCGGAC	1440
Qy	1441	AAATCTCCCAATTCACGGAGAAATCCCAAGAGGAGTTTGGAAATAAAGGCTTAACAGA	1500
Db	1441	AAATCTCCCAATTCACGGAGAAATCCCAAGAGGAGTTTGGAAATAAAGGCTTAACAGA	1500
Qy	1501	AAAGTGTTTGGCCAGTGGGTGTCCAGCAGTACTCTAAATATGATGAGGCGCGGAAAGGG	1560
Db	1501	AAAGTGTTTGGCCAGTGGGTGTCCAGCAGTACTCTAAATATGATGAGGCGCGGAAAGGG	1560
Qy	1561	CTTCTGGTCAAATGCTGTGCGTACAAATCTATGAGGACAGTCTTCGGAACCATCTTTTAA	1620
Db	1561	CTTCTGGTCAAATGCTGTGCGTACAAATCTATGAGGACAGTCTTCGGAACCATCTTTTAA	1620
Qy	1621	AACTCAGAGGCCCTTCCCTTTTCATCCAAGACACATCTCTCTATGCGACCATTTGGG	1680
Db	1621	AACTCAGAGGCCCTTCCCTTTTCATCCAAGACACATCTCTCTATGCGACCATTTGGG	1680
Qy	1681	CTCTGCTCTCCCTTCATTTGTTGTTCTCTATGTTGTTGATCTGCCACAAATACAAAAGCAA	1740
Db	1681	CTCTGCTCTCCCTTCATTTGTTGTTCTCTATGTTGTTGATCTGCCACAAATACAAAAGCAA	1740
Qy	1741	TTTAGGTACGAGAGTCAGCTGCAGATGATCCAGGTGACTGCCCCCTTGGATTAACGAGTAC	1800
Db	1741	TTTAGGTACGAGAGTCAGCTGCAGATGATCCAGGTGACTGCCCCCTTGGATTAACGAGTAC	1800
Qy	1801	TTTCTACGTTGACTTCAGGACATATGAATATAGCTTAAGTGGGAGTTCCCGCAGAGAGAAC	1860

[illegible]

Db 2881 GCAGAGGAGAGAGCATGTATCAGAACATCCATCATCTACCAAAACAGCGGCCCT 2940
 QY 2941 CAGCAGAGAGCGGCTCAGAGCCAGCTCGCCACAGCGCCAGGTGAAGATTCAAGAGAA 3000
 Db 2941 CAGCAGAGAGCGGCTCAGAGCCAGCTCGCCACAGCGCCAGGTGAAGATTCAAGAGAA 3000
 QY 3001 AGAAGTTAGGAGAGAGCGCTTGGACCCCGCCACCTAGCAGGCTGTAGCGCAGAGCCA 3060
 Db 3001 AGAAGTTAGGAGAGAGCGCTTGGACCCCGCCACCTAGCAGGCTGTAGCGCAGAGCCA 3060
 QY 3061 AGATTAGCCTCGCTCTGAGGAAGCGCCCTACAGCGGTTGCTTCGCTGGACTTTTCTCT 3120
 Db 3061 AGATTAGCCTCGCTCTGAGGAAGCGCCCTACAGCGGTTGCTTCGCTGGACTTTTCTCT 3120
 QY 3121 AGATGCTCTGCTGCATCTACCTCAAGTACTCTATAAATCAAACTCTCTCGCACAG 3180
 Db 3121 AGATGCTCTGCTGCATCTACCTCAAGTACTCTATAAATCAAACTCTCTCGCACAG 3180
 QY 3181 GCGGAGAGGCAATAATGAGACTTGTGGTGAGCGCCCTACCTCGGGGCTTTCACG 3240
 Db 3181 GCGGAGAGGCAATAATGAGACTTGTGGTGAGCGCCCTACCTCGGGGCTTTCACG 3240
 QY 3241 AGCTTGAAGGAAAGCCATGTATCTGAAATATAGTATATCTTGTAAATACGTGAACAA 3300
 Db 3241 AGCTTGAAGGAAAGCCATGTATCTGAAATATAGTATATCTTGTAAATACGTGAACAA 3300
 QY 3301 ACCAAACCCGTTTTTGTCTAAGGAAAGCTAAATATGATTTTAAATCTATGTTTTAA 3360
 Db 3301 ACCAAACCCGTTTTTGTCTAAGGAAAGCTAAATATGATTTTAAATCTATGTTTTAA 3360
 QY 3361 AATACTATGTAATTTTTCATCTATTATGATATATTTATGATGATATTTATGATGATGAT 3420
 Db 3361 AATACTATGTAATTTTTCATCTATTATGATATATTTATGATGATATTTATGATGATGAT 3420
 QY 3421 TACTGTAAAAAATAAAAAAAAAAAAAAAAAAAAAA 3453
 Db 3421 TACTGTAAAAAATAAAAAAAAAAAAAAAAAAAAAA 3453

RESULT 6

AAQ53502
 ID AAQ53502 standard; cDNA; 3453 BP.
 XX
 AC AAQ53502;
 XX
 DT 25-MAR-2003 (updated)
 DT 27-JUN-1994 (first entry)
 XX
 DE Murine flk-2 cDNA.
 XX
 KW Receptor protein tyrosine kinase; ptk family; foetal liver kinase;
 KW mflk; primitive; totipotent; haematopoietic cell; stem cell;
 KW proliferation; stromal cell; ds.
 XX
 OS Mus musculus.
 XX
 FH Key Location/Qualifiers
 FT CDS 31..3009
 FT /*tag= a
 FT sig_peptide 31..111
 FT /*tag= b
 FT mat_peptide 112..3006
 FT /*tag= c
 XX
 FN US5270458-A.
 XX
 PD 14-DEC-1993.
 XX
 PE 19-NOV-1992; 92US-0977451.
 XX
 PR 02-APR-1991; 91US-0679666.
 PR 28-JUN-1991; 91US-0728913.
 PR 15-NOV-1991; 91US-0793065.

PR 24-DEC-1991; 91US-0813593.
 PR 26-JUN-1992; 92US-0906397.
 PR 12-NOV-1992; 92US-0975049.
 PR 19-NOV-1992; 92US-0977451.
 XX
 PA (UUPR-) UNIV PRINCETON.
 XX
 XX Lemischka IR;
 XX
 XX WPI; 1993-405021/50.
 DR P-PSDB; AAR44994.
 XX
 XX Isolated nucleic acid molecules of hematopoietic stem cell
 PT receptor flk-2 - encoding mammalian receptor protein tyrosine
 PT kinases expressed in primitive haematopoietic cells
 XX
 PS Claim 2; Fig 1a; 60pp; English.
 XX
 XX Nucleic acid sequences coding for murine flk-2 and specified
 CC subfragments of it are claimed. The flk-2 polypeptide is a protein
 CC tyrosine kinase expressed only in primitive haematopoietic cells.
 CC The cDNA can be used to recombinantly produce flk-2 for stimulating
 CC self-renewal of totipotent stem cells and development of all
 CC haematopoietic cells.
 CC (Updated on 25-MAR-2003 to correct PF field.)
 XX
 SQ Sequence 3453 BP; 947 A; 821 C; 876 G; 809 T; 0 other;
 Query Match 100.0%; Score 3451.4; DB 14; Length 3453;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 3452; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 GCGGCTGCTGCTACCGCGCTCGGAGGCAATCGCGGCGTTCGCGAGCGACGACCGG 60
 Db 1 GCGGCTGCTGCTACCGCGCTCGGAGGCAATCGCGGCGTTCGCGAGCGACGACCGG 60
 QY 61 CGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
 Db 61 CGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
 QY 121 CTGCTGCTGATCAAGTGTGTTTTTAATCATGTCATGAGAACATGGCTCATCAGCGGAAAG 180
 Db 121 CTGCTGCTGATCAAGTGTGTTTTTAATCATGTCATGAGAACATGGCTCATCAGCGGAAAG 180
 QY 181 CCATCATGTCACCGAATGCTGCGAGGATCCCGAGAGCTCCAGAGACCTCAGTGTACCGAGCGC 240
 Db 181 CCATCATGTCACCGAATGCTGCGAGGATCCCGAGAGACCTCCAGTGTACCGAGCGC 240
 QY 241 CAGAGTGAAGGACCGGTATATGAAGCGGCCACCGTGGAGGTGGCGAGTCTGGTCCATC 300
 Db 241 CAGAGTGAAGGACCGGTATATGAAGCGGCCACCGTGGAGGTGGCGAGTCTGGTCCATC 300
 QY 301 ACCCTGCAAGTGACAGTCTGCGCCACCCAGGAGACCTTCTGCTCTGGGTCTTTAAGCAC 360
 Db 301 ACCCTGCAAGTGACAGTCTGCGCCACCCAGGAGACCTTCTGCTCTGGGTCTTTAAGCAC 360
 QY 361 AGCTCCCTGGGCTGCGAGCGCACTTTGATTACAAACAGAGGAATCGTTTCCATGGCC 420
 Db 361 AGCTCCCTGGGCTGCGAGCGCACTTTGATTACAAACAGAGGAATCGTTTCCATGGCC 420
 QY 421 ATCTTGAACGTGACAGAGACCCAGCAGGAGAAATACCTACTCCATATTCAGAGCGAAACG 480
 Db 421 ATCTTGAACGTGACAGAGACCCAGCAGGAGAAATACCTACTCCATATTCAGAGCGAAACG 480
 QY 481 GCCAACTACACAGTACTGTTTACAGTGAATGTAAGAGATACACAGCTGTATGTCTAAGG 540
 Db 481 GCCAACTACACAGTACTGTTTACAGTGAATGTAAGAGATACACAGCTGTATGTCTAAGG 540
 QY 541 AGACCTTACTTTAGGAAGATGGAACACGAGATGCACTGTCTGCATCTCCGAGGGTGT 600
 Db 541 AGACCTTACTTTAGGAAGATGGAACACGAGATGCACTGTCTGCATCTCCGAGGGTGT 600
 QY 601 CCGAGGCCACTGTGGAGTGGGTGCTCTGAGCTCCACAGGGAAGCTGTAAAGAGAA 660

601	CGGAGCCACTGTGAGTGGGTCTCTGAGCTCCACAGGGAAGCTGTAAAGAGAA	660	1681	CTCTGTCTCCCTTCAATGTTGTTCTCAATGTTGTTGATCTGCCACAATAACAAAAGCAA	1740
661	GGCCCTGCTGTTGTGAGAAAGGGAAGAGTACTTTCATGAGTTGTTTCGGAACAGACATC	720	1741	TTTAGTACAGAGTCAAGTGTGAGATGATCCAGTGTGATCTGCCCCCTGGATACAGAGTAC	1800
661	GGCCCTGCTGTTGTGAGAAAGGGAAGAGTACTTTCATGAGTTGTTTCGGAACAGACATC	720	1741	TTTAGTACAGAGTCAAGTGTGAGATGATCCAGTGTGATCTGCCCCCTGGATACAGAGTAC	1800
721	AGATGCTGTGTAGAAATGCACTGGGCGCGGAATGCCAACAGCTGTTTCAACATAGATCTA	780	1801	TTCTACGTTGACTTCAGGGACTATGAATATGACCTTAAAGTGGGAGTTCCCGAGAGAGAAC	1860
721	AGATGCTGTGTAGAAATGCACTGGGCGCGGAATGCCAACAGCTGTTTCAACATAGATCTA	780	1801	TTCTACGTTGACTTCAGGGACTATGAATATGACCTTAAAGTGGGAGTTCCCGAGAGAGAAC	1860
781	AACACAGGCTCTCAGAGCACTGCCCCAGTTATCTTGAAGTGGGGGAACCTTTGTGG	840	1861	TTAGAGTTTGGGAAGTCTGCGGCTTTCGGGAGGGTGTGAACGCCACCGGCC	1920
781	AACACAGGCTCTCAGAGCACTGCCCCAGTTATCTTGAAGTGGGGGAACCTTTGTGG	840	1861	TTAGAGTTTGGGAAGTCTGCGGCTTTCGGGAGGGTGTGAACGCCACCGGCC	1920
841	ATCAGTGTAAAGCCATCATGTGAACCATGGATTCGGGCTCACTGGGAGCTGGAAGAC	900	1921	TATGGCATTTAGTAAAACGGGAGTCTCAATTCAGTGGCGGTGAAGATGCTAAAAGAGAAA	1980
841	ATCAGTGTAAAGCCATCATGTGAACCATGGATTCGGGCTCACTGGGAGCTGGAAGAC	900	1921	TATGGCATTTAGTAAAACGGGAGTCTCAATTCAGTGGCGGTGAAGATGCTAAAAGAGAAA	1980
901	AAAGCCCTGGAGGAGGAGCTACTTTGAGATGAGTACCTTCAACAAACAGGACCATG	960	1981	GCTGACAGCTGTGAAAAGAAAGCTCTCATGTGCGAGCTCAAAATGATGACCCACCTGGGA	2040
901	AAAGCCCTGGAGGAGGAGCTACTTTGAGATGAGTACCTTCAACAAACAGGACCATG	960	1981	GCTGACAGCTGTGAAAAGAAAGCTCTCATGTGCGAGCTCAAAATGATGACCCACCTGGGA	2040
961	ATTTCGGATTCTTTGGCTTTGTCTTCCGTGGGAAGGAACGACACCGGATATTACAC	1020	2041	CACCATGACAAACATCGTGAATCTGCTGGGGCATGCAACATGTCAGGGCCAGTGTACTTG	2100
961	ATTTCGGATTCTTTGGCTTTGTCTTCCGTGGGAAGGAACGACACCGGATATTACAC	1020	2041	CACCATGACAAACATCGTGAATCTGCTGGGGCATGCAACATGTCAGGGCCAGTGTACTTG	2100
1021	TGCTCTTCCTCAAGACACCCAGCAGTACCGTTGGTGACCATCTTAGAAAAGGGTTT	1080	2101	ATTTTGAATATTGTTGCTATGTTGCTCTCTCAACTACCTAAAGAGTAAAGAGAGAG	2160
1021	TGCTCTTCCTCAAGACACCCAGCAGTACCGTTGGTGACCATCTTAGAAAAGGGTTT	1080	2101	ATTTTGAATATTGTTGCTATGTTGCTCTCTCAACTACCTAAAGAGTAAAGAGAGAG	2160
1081	ATAACGCTACCACTCGCAGAGAGTATGAATTTGACCCGTACGAAAGTCTTGCTTC	1140	2161	TTTTCACAGGACATGACAGAGATTTTAAAGAAACATAATTTTCACTTACCTACTTTTC	2220
1081	ATAACGCTACCACTCGCAGAGAGTATGAATTTGACCCGTACGAAAGTCTTGCTTC	1140	2161	TTTTCACAGGACATGACAGAGATTTTAAAGAAACATAATTTTCACTTACCTACTTTTC	2220
1141	TCAGTCAAGTTTAAAGCGTACCCAGCAATCCGATCCGATGCGATCTTCTCAAGCCCTCA	1200	2221	CAGGACATTCAAATTCAGCATGCTGTTTCAAGAGAGTTCAGTTACACCCGCCCTTG	2280
1141	TCAGTCAAGTTTAAAGCGTACCCAGCAATCCGATCCGATGCGATCTTCTCAAGCCCTCA	1200	2221	CAGGACATTCAAATTCAGCATGCTGTTTCAAGAGAGTTCAGTTACACCCGCCCTTG	2280
1201	TTTCTTTGTAACAGAGGCTGGAGTGGGTACAGCATATCTAAATTTTGGGATCAT	1260	2281	GATCAGCTCTCAGGGTTCAATTCAGGAAATTCATTCCTGAAGATGAGATTAATGAA	2340
1201	TTTCTTTGTAACAGAGGCTGGAGTGGGTACAGCATATCTAAATTTTGGGATCAT	1260	2281	GATCAGCTCTCAGGGTTCAATTCAGGAAATTCATTCCTGAAGATGAGATTAATGAA	2340
1261	AAGAACAGCCAGAGAGTACATATCTATCAGAAATGATGACCCCGTACACCAAA	1320	2341	AACAGAGAGCTGGCAG	2400
1261	AAGAACAGCCAGAGAGTACATATCTATCAGAAATGATGACCCCGTACACCAAA	1320	2341	AACAGAGAGCTGGCAG	2400
1321	ATGTTCAAGCTGAATATAAGAAAGAAACCTCAAGTGTAGCAAAATGCTCAGCCAGCAG	1380	2401	CTCTCTTCTTTGCTTACCAAGTGGCCAAAGGATGGAATTCCTCGAGTTCAAGTCTGT	2460
1321	ATGTTCAAGCTGAATATAAGAAAGAAACCTCAAGTGTAGCAAAATGCTCAGCCAGCAG	1380	2401	CTCTCTTCTTTGCTTACCAAGTGGCCAAAGGATGGAATTCCTCGAGTTCAAGTCTGT	2460
1381	CGCTCTGTTCTCTGATGGCTACCCGCTACCTCTTGGACCTGGAAGAGTGTTCGGAC	1440	2461	GTCCAACAGAGACCTGGCAGCCAGGAATGTGTTGGTCAACCAAGAGAGTGTGAAGATC	2520
1381	CGCTCTGTTCTCTGATGGCTACCCGCTACCTCTTGGACCTGGAAGAGTGTTCGGAC	1440	2461	GTCCAACAGAGACCTGGCAGCCAGGAATGTGTTGGTCAACCAAGAGAGTGTGAAGATC	2520
1441	AAATCTCCAAATGACCGAGGAATCCAGAGAGTGTGGAATATAAGAGCTTAACAGA	1500	2521	TGTGACTTTGACTGGCCCGGAGACATCTGAGCGACTCCAGTACGTCGTCAGGGGCAAC	2580
1441	AAATCTCCAAATGACCGAGGAATCCAGAGAGTGTGGAATATAAGAGCTTAACAGA	1500	2521	TGTGACTTTGACTGGCCCGGAGACATCTGAGCGACTCCAGTACGTCGTCAGGGGCAAC	2580
1501	AAAGTGTGTCAGTGGTGTGAGAGCACTCTTAAATATGATGAGTGGCCGGGAAAGGG	1560	2581	GCACGGCTGCGGCTCAAGTGGATGCAACCCGAGAGCTTATTTGAAGGGATCTACAAATC	2640
1501	AAAGTGTGTCAGTGGTGTGAGAGCACTCTTAAATATGATGAGTGGCCGGGAAAGGG	1560	2581	GCACGGCTGCGGCTCAAGTGGATGCAACCCGAGAGCTTATTTGAAGGGATCTACAAATC	2640
1561	CTTCTGGTCAAAATGCTGCTGACAAATCTATGGGCACTGTTGGGAACCAATCTTTTAA	1620	2641	AAGAGTGAAGTCTGCTTACCGCATCTTCTCTGGAGATCTTCTCTGGGATATTTTCACTGGGTGAAC	2700
1561	CTTCTGGTCAAAATGCTGCTGACAAATCTATGGGCACTGTTGGGAACCAATCTTTTAA	1620	2641	AAGAGTGAAGTCTGCTTACCGCATCTTCTCTGGAGATCTTCTCTGGGATATTTTCACTGGGTGAAC	2700
1621	AACCTACAGGCCCCCTTCCCTTTCATCCAGACCAATCTCTTATGCGACCAATGGG	1680	2701	CCTTACCTTGGCAATCTCTGTCGACGCTAACTTCTATATAAAGTCTGATTCAGAGTGAATATAA	2760
1621	AACCTACAGGCCCCCTTCCCTTTCATCCAGACCAATCTCTTATGCGACCAATGGG	1680	2701	CCTTACCTTGGCAATCTCTGTCGACGCTAACTTCTATATAAAGTCTGATTCAGAGTGAATATAA	2760
1681	CTCTGTCTCCCTTCAATGTTGTTCTCAATGTTGTTGATCTGCCACAATAACAAAAGCAA	1740	2761	ATGGAGCAGCCATCTTATGCCAGAGGAGATATCTTTTGAATGCAATCTCTGCTGGGCT	2820

QY	2821	TTTGACTCAAGGAAGCGGCATCTTCTCCCAACCTGACTTTCATTTTATAGGATGTCAGCTG	2880
Db	2821	TTTGACTCAAGGAAGCGGCATCTTCCCAACCTGACTTTCATTTTATAGGATGTCAGCTG	2880
QY	2881	GCAGAGGCAGAGAAGCATGTATCAGAACATCCATCCATCTACTACAAACACAGCGCGCCCT	2940
Db	2881	GCAGAGGCAGAGAAGCATGTATCAGAACATCCATCCATCTACTACAAACACAGCGCGCCCT	2940
QY	2941	CAGCAGAGAGCGCGCTCAGAGCCACGTCGCCACAGCGCCAGGTGCAAGTTCACAGAGAA	3000
Db	2941	CAGCAGAGAGCGCGCTCAGAGCCACGTCGCCACAGCGCCAGGTGCAAGTTCACAGAGAA	3000
QY	3001	AGAAAGTTACGAGAGGCGCTTGGAACCCCGCCACCCCTAGCAGCTGTAGACCGCAGAGCCA	3060
Db	3001	AGAAAGTTACGAGAGGCGCTTGGAACCCCGCCACCCCTAGCAGCTGTAGACCGCAGAGCCA	3060
QY	3061	AGATTAGCTCGCCTCTCAGGAGCGCCCTACAGCGCGTTCCTTCGCTGGACTTTTCTCT	3120
Db	3061	AGATTAGCTCGCCTCTCAGGAGCGCCCTACAGCGCGTTCCTTCGCTGGACTTTTCTCT	3120
QY	3121	AGATGCTGTCTGCCATTACTCCAAAGTCATCTCTATAAATCAAACTCTCTCTCGCACAG	3180
Db	3121	AGATGCTGTCTGCCATTACTCCAAAGTCATCTCTATAAATCAAACTCTCTCTCGCACAG	3180
QY	3181	GGGGAGAGCCCAATTAATGAGACTTGTGTGGTGAGCGCGCTACCCCTGGGGCGCTTTCACAG	3240
Db	3181	GGGGAGAGCCCAATTAATGAGACTTGTGTGGTGAGCGCGCTACCCCTGGGGCGCTTTCACAG	3240
QY	3241	AGCTTGAGGGGAAAGCCCATGTATCTGAAATAGTATATTTCTGTAAATACGTGAAACAA	3300
Db	3241	AGCTTGAGGGGAAAGCCCATGTATCTGAAATAGTATATTTCTGTAAATACGTGAAACAA	3300
QY	3301	ACCAAAACCGTTTTTTTGTCTAGGGAAGCTAAATATCATTTTTTAAAAATCTATGTTTTAA	3360
Db	3301	ACCAAAACCGTTTTTTTGTCTAGGGAAGCTAAATATCATTTTTTAAAAATCTATGTTTTAA	3360
QY	3361	AATACTATGTAACTTTTTCAPTATTTAGTGATATATTTTATGATGGAATAAACTTTTC	3420
Db	3361	AATACTATGTAACTTTTTCAPTATTTAGTGATATATTTTATGATGGAATAAACTTTTC	3420
QY	3421	TACTGTAAAAAATAAAAAAAAAAAAAAAAAA	3453
Db	3421	TACTGTAAAAAATAAAAAAAAAAAAAAAAAA	3453
RESULT	7		
AAQ35249			
ID	AAQ35249	standard; cdna; 3453 BP.	
XX	XX		
AC	AAQ35249;		
XX	XX		
DT	25-MAR-2003	(updated)	
DT	25-JUN-1993	(first entry)	
XX	XX		
DE	Murine flk-2 coding sequence.		
XX	XX		
KW	Murine; receptor; protein; tyrosine kinase; ptk; primitive; mammalian;		
KW	hematopoietic cell; phc; mature; mhc; fetal; liver kinase 2; flk-2;		
KW	liver; spleen; thymus; adult; brain; marrow; thymocyte; subset;		
KW	multipotential; T-lymphoid; lineage; ss.		
XX	XX		
OS	Mus musculus.		
XX	XX		
FH	Key	Location/Qualifiers	
CDS		31..3009	
FT	FT	/*tag= a	
FT	FT	31..111	
FT	FT	/*tag= b	
FT	FT	/*tag= "Hydrophobic leader"	
XX	XX		
PN	WO9300349-A1.		
XX	XX		

481 GCCAATACACAGTACTGTTACAGTGAATGTAGAGATACACAGCTGTATGTCTAAGG 540
481 GCCAATACACAGTACTGTTACAGTGAATGTAGAGATACACAGCTGTATGTCTAAGG 540
541 AGACCTTACTTTAGGAAGATGGAACACAGGATGCACTGCTCTGCACTCTCCAGGGTGT 600
541 AGACCTTACTTTAGGAAGATGGAACACAGGATGCACTGCTCTGCACTCTCCAGGGTGT 600
601 CCGAGCCCACTGTGGAGTGGTCTCTGAGCTCCACAGGAAAGCTGTAAAGAGAA 660
601 CCGAGCCCACTGTGGAGTGGTCTCTGAGCTCCACAGGAAAGCTGTAAAGAGAA 660
661 GGCCCTGCTGTTCAGAGAGGAAAGTACTTCTATGATGTTGTTCCGGAACAGACATC 720
661 GGCCCTGCTGTTCAGAGAGGAAAGTACTTCTATGATGTTGTTCCGGAACAGACATC 720
721 AGATGCTGTCTAGAAATGCACTGGGCGCGAATGCACCAAGCTGTTCCACATAGATCTA 780
721 AGATGCTGTCTAGAAATGCACTGGGCGCGAATGCACCAAGCTGTTCCACATAGATCTA 780
781 AACAGGCTCTCTAGAGACACTGCCCCAGTTATTTCTGAAAGTGGGGAAACCTTGTGG 840
781 AACAGGCTCTCTAGAGACACTGCCCCAGTTATTTCTGAAAGTGGGGAAACCTTGTGG 840
841 ATCAGGTGTAAAGGCATCCATGTGAACCATGGATTCGGGCTCACCTGGAGCTGGAAGC 900
841 ATCAGGTGTAAAGGCATCCATGTGAACCATGGATTCGGGCTCACCTGGAGCTGGAAGC 900
901 AAAGCCCTGGAGGGGAGCTACTTTGAGATGAGTACTCTCCACAAACAGGACCATG 960
901 AAAGCCCTGGAGGGGAGCTACTTTGAGATGAGTACTCTCCACAAACAGGACCATG 960
961 ATTCCGATTTCTTTGGCCCTTTGTCTTCCGTTGGGAAGAACACACCGGATATTACACC 1020
961 ATTCCGATTTCTTTGGCCCTTTGTGTCTTCCGTTGGGAAGAACACACCGGATATTACACC 1020
1021 TGCTCTTCTCAAGCACCACAGCAGTCTGAGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1080
1021 TGCTCTTCTCAAGCACCACAGCAGTCTGAGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1080
1081 ATAAAGCTTACAGCTCCGAGAGAGATGAAATGGACCTGACGAAAGTTCTGCTTC 1140
1081 ATAAAGCTTACAGCTCCGAGAGAGATGAAATGGACCTGACGAAAGTTCTGCTTC 1140
1141 TCAGTCAGGTTTAAAGCGTACCACGAATCCGATGCACTGGATCTTCTCAAGCCTCA 1200
1141 TCAGTCAGGTTTAAAGCGTACCACGAATCCGATGCACTGGATCTTCTCAAGCCTCA 1200
1201 TTTCTTGTGAAACAGAGAGGCTGGAGGATGGGTACAGATATCTTAAATTTTGGCATCAT 1260
1201 TTTCTTGTGAAACAGAGAGGCTGGAGGATGGGTACAGATATCTTAAATTTTGGCATCAT 1260
1261 AAGAACAGCAGGAGAGTACATATCTATGATGAGAAATGAGTACGCTTACCAAA 1320
1261 AAGAACAGCAGGAGAGTACATATCTATGATGAGAAATGAGTACGCTTACCAAA 1320
1321 ATGTTTCAGCTCAATATAGAAAGAACCTCAAGTGTAGCAATGCTTACGAGCCAG 1380
1321 ATGTTTCAGCTCAATATAGAAAGAACCTCAAGTGTAGCAATGCTTACGAGCCAG 1380
1381 GCGTCTGTTCTCTGATGGCTACCGCTACCCCTTTGAGCCTGGAGAGTGTTCGGAC 1440
1381 GCGTCTGTTCTCTGATGGCTACCGCTACCCCTTTGAGCCTGGAGAGTGTTCGGAC 1440
1441 AAATCTCCAAATGCAACGAGGAAATCCAGAGAGGATTTGGATATAAAGGCTACAGA 1500
1441 AAATCTCCAAATGCAACGAGGAAATCCAGAGAGGATTTGGATATAAAGGCTACAGA 1500
1501 AAAGTGTGTTGGCAGTGGGTGTCGAGCAGTACTCTAAATATAGTACGAGCCGGAAGGG 1560
1501 AAAGTGTGTTGGCAGTGGGTGTCGAGCAGTACTCTAAATATAGTACGAGCCGGAAGGG 1560
1561 CTTCTGCTCAAAATGCTGTGCGTACAAATCTATGGGCAAGTCTTTGCGAAACCATCTTTT 1620

1561 CTTCTGCTCAAAATGCTGTGCGTACAAATCTATGGGCAAGTCTTGGAAACCATCTTTT 1620
1621 AACTCACAGGCCCCCTTCCCTTTTCATCAAGACAACATCTCTCTATGCGACCATTTGG 1680
1621 AACTCACAGGCCCCCTTCCCTTTTCATCAAGACAACATCTCTCTATGCGACCATTTGG 1680
1681 CTTCTGCTCCCCCTTCAATTTGTTCTCTCAATTTGTTGATCTGCTCAAAATACAAAAGCAA 1740
1681 CTTCTGCTCCCCCTTCAATTTGTTCTCTCAATTTGTTGATCTGCTCAAAATACAAAAGCAA 1740
1741 TTTAGGTACGAGATGAGTCCAGATGATCCAGTGTACTGGCCCCCTGGATACAGGTAC 1800
1741 TTTAGGTACGAGATGAGTCCAGATGATCCAGTGTACTGGCCCCCTGGATACAGGTAC 1800
1801 TTTACGTTGAGTTCAGGGACTATGAATATGACCTTAAAGTGGAGTTCCCGAGAGAGAAC 1860
1801 TTTACGTTGAGTTCAGGGACTATGAATATGACCTTAAAGTGGAGTTCCCGAGAGAGAAC 1860
1861 TTAGAGTTTGGGAAGTCTCTGGGGTCTGGCGCTTTCGGGAGGGTGAACCCACGGCC 1920
1861 TTAGAGTTTGGGAAGTCTCTGGGGTCTGGCGCTTTCGGGAGGGTGAACCCACGGCC 1920
1921 TATGSCATTAGTAAACCGGAGTCTCAATTCAGTGTGGCGTCAAGATGCTAAAGAGAA 1980
1921 TATGSCATTAGTAAACCGGAGTCTCAATTCAGTGTGGCGTCAAGATGCTAAAGAGAA 1980
1981 GCTGACAGCTGTGAAAAAGAGCTCTCATGTCGGAGCTCAAAATGATGACCCACCTGGGA 2040
1981 GCTGACAGCTGTGAAAAAGAGCTCTCATGTCGGAGCTCAAAATGATGACCCACCTGGGA 2040
2041 CACCATGACACATCGTGAATCTGCTGGGGCATGCACACTCTCAGGGCCAGTGTACTTG 2100
2041 CACCATGACACATCGTGAATCTGCTGGGGCATGCACACTCTCAGGGCCAGTGTACTTG 2100
2101 ATTTTGAATATTTGTTGCTATGCTGACCTCTCAACTACCTAAGAGTAAAGAGAGAG 2160
2101 ATTTTGAATATTTGTTGCTATGCTGACCTCTCAACTACCTAAGAGTAAAGAGAGAG 2160
2161 TTTTCAAGACATGACAGAGATTTTAAAGAACATAATTTCACTTCTTACCTTCTTTC 2220
2161 TTTTCAAGACATGACAGAGATTTTAAAGAACATAATTTCACTTCTTACCTTCTTTC 2220
2221 CAGGCACATTCCTCAATTCAGCATGCTGCTGCTCAGAGAGTTCAGTTACCCGCCCTTG 2280
2221 CAGGCACATTCCTCAATTCAGCATGCTGCTGCTCAGAGAGTTCAGTTACCCGCCCTTG 2280
2281 GATCAGCTCTCAGGGTTCATGCGAATTCATTTCAATTTCAATTTCAATTTCAATTTCA 2340
2281 GATCAGCTCTCAGGGTTCATGCGAATTCATTTCAATTTCAATTTCAATTTCAATTTCA 2340
2341 AACCCAGAGAGCTGGCAGAGAGAGAGAGATTTGAAACGCTGCTGAGCTTTGAAGAC 2400
2341 AACCCAGAGAGCTGGCAGAGAGAGAGAGATTTGAAACGCTGCTGAGCTTTGAAGAC 2400
2401 CTTCTTTGTTTTCGTTACCAAGTGGCCAAAGGATGCTGCTGAGTTCAAGTCTGCT 2460
2401 CTTCTTTGTTTTCGTTACCAAGTGGCCAAAGGATGCTGCTGAGTTCAAGTCTGCT 2460
2461 GTTCCACAGAGAGCTGGCAGAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2520
2461 GTTCCACAGAGAGCTGGCAGAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2520
2521 TGTGATTTTGGACTGGCCCGGAGATCTCTGAGCGATCTCCAGCTACGTCGTGAGGGCAAC 2580
2521 TGTGATTTTGGACTGGCCCGGAGATCTCTGAGCGATCTCCAGCTACGTCGTGAGGGCAAC 2580
2581 GCACGCTGCTGGTGAAGTGGATGCGACCCGAGAGCTTATTTGAAGGGATCTACACATC 2640
2581 GCACGCTGCTGGTGAAGTGGATGCGACCCGAGAGCTTATTTGAAGGGATCTACACATC 2640
2641 AAGAGTGAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2700

```

Db 2641 AAGAGTACGCTCTGGTCTCTAGCGCATCCTTCTCTGGGAGATATTTTCACTGGGTGTGAAC 2700
Qy 2701 CCTTACCTGGCAFTTCTGTGCGAGCTAACTTCTATAAACTGATTCAGAGTGGATTTAAA 2760
Db 2701 CCTTACCTGGCAFTTCTGTGCGAGCTAACTTCTATAAACTGATTCAGAGTGGATTTAAA 2760
Qy 2761 ATGAGCAGCATCTTATGACACAGAGGATATATTTGTATGCAATCTCTGCTGGCT 2820
Db 2761 ATGAGCAGCATCTTATGACACAGAGGATATATTTGTATGCAATCTCTGCTGGCT 2820
Qy 2821 TTTGACTCAAGGAAGCGGCCTCTTCCCAACTGACTTCATTTTAGGATGTCAGCTG 2880
Db 2821 TTTGACTCAAGGAAGCGGCCTCTTCCCAACTGACTTCATTTTAGGATGTCAGCTG 2880
Qy 2881 GCAGAGCAGAGAGCATGTATCAGAACTATCCATCCATCTACCAAAACAGGGCGCCCT 2940
Db 2881 GCAGAGCAGAGAGCATGTATCAGAACTATCCATCCATCTACCAAAACAGGGCGCCCT 2940
Qy 2941 CAGCAGAGGGGGCTCAGAGCCAGTCCGACAGCGCCAGGTGAAGATTACAGAGAA 3000
Db 2941 CAGCAGAGGGGGCTCAGAGCCAGTCCGACAGCGCCAGGTGAAGATTACAGAGAA 3000
Qy 3001 AGAAGTTAGCGAGGAGCGCTTGAGCCCGCCAGCCCTAGCAGGCTGTAGCCGAGAGCCA 3060
Db 3001 AGAAGTTAGCGAGGAGCGCTTGAGCCCGCCAGCCCTAGCAGGCTGTAGCCGAGAGCCA 3060
Qy 3061 AGATTAGCTCGCTCTGAGGAAGCGCCCTCAGAGCGCTTCTGCTGGACTTTCTCT 3120
Db 3061 AGATTAGCTCGCTCTGAGGAAGCGCCCTCAGAGCGCTTCTGCTGGACTTTCTCT 3120
Qy 3121 AGATCTGTCTGCCATTACTCCAAAGTGACTTCTATAAATCAAACTCTCTCGCACAG 3180
Db 3121 AGATCTGTCTGCCATTACTCCAAAGTGACTTCTATAAATCAAACTCTCTCGCACAG 3180
Qy 3181 GCGGAGAGCCATATATGAGACTTGTGTGAGCGCCCTACCTGGGGCGCTTTCCAGC 3240
Db 3181 GCGGAGAGCCATATATGAGACTTGTGTGAGCGCCCTACCTGGGGCGCTTTCCAGC 3240
Qy 3241 AGCTTGAGGGGAAAGCCATGTATCTGAAATAGTATATTTCTGTAATACGTGAACAA 3300
Db 3241 AGCTTGAGGGGAAAGCCATGTATCTGAAATAGTATATTTCTGTAATACGTGAACAA 3300
Qy 3301 ACCAAACCCGTTTTTGTAAAGGAAAGCTAAATATGATTTTAAATCTATGTTTTAA 3360
Db 3301 ACCAAACCCGTTTTTGTAAAGGAAAGCTAAATATGATTTTAAATCTATGTTTTAA 3360
Qy 3361 AATACTATGTAACCTTTTCTATCTATTTAGTGATATATTTTATGATGAAATAAACTTTC 3420
Db 3361 AATACTATGTAACCTTTTCTATCTATTTAGTGATATATTTTATGATGAAATAAACTTTC 3420
Qy 3421 TACTGTAAAAAATAAAAAAAAAAAAAAAAAAAAAA 3453
Db 3421 TACTGTAAAAAATAAAAAAAAAAAAAAAAAAAAAA 3453

```

RESULT 8

```

AAQ79068
ID AAQ79068 standard; cDNA; 3453 BP.
XX
AC AAQ79068;
XX
XX
DT 25-MAR-2003 (updated)
DT 04-JUL-1995 (first entry)
XX
DE Mouse flk-2 cDNA.
XX
KW Fetal liver kinase-2; flk-2; protein tyrosine-kinase receptor;
KW hematopoiesis; stem cell; ds.
XX
OS Mus sp.
XX
FH Key
CDS 58..3039

```

```

FT sig_peptide /*tag= a
FT 58..138 /*tag= b
FT mat_peptide 139..3036 /*tag= c
FT
XX US5367057-A.
XX
XX 22-NOV-1994.
XX
XX 30-APR-1993; 93US-0055269.
XX
XX 02-APR-1991; 91US-0679666.
XX 28-JUN-1991; 91US-0728913.
XX 15-NOV-1991; 91US-0793065.
XX 24-DEC-1991; 91US-0813593.
XX 26-JUN-1992; 92US-0906397.
XX 12-NOV-1992; 92US-0975049.
XX 19-NOV-1992; 92US-0977451.
XX 30-APR-1993; 93US-0055269.
XX
XX (UYPR-) UNIV PRINCETON.
XX
XX Lemischka IR;
XX
XX WPI; 1995-005894/01.
XX P-PSDB; AAR67535.
XX
XX Murine flk-2 receptor protein tyrosine kinase - used to stimulate
XX proliferation and/or stimulation of primitive mammalian
XX hematopoietic stem cells in vitro or in vivo.
XX
XX Disclosure; Fig. 1A-1F; 69pp; English.
XX
XX cDNAs encoding receptor protein tyrosine-kinases, mouse fetal liver
XX kinase-2 (flk-2), human flk-2 and mouse flk-1 are given in AAQ79068-70,
XX respectively, and the deduced amino acid sequences in AAR67535-37,
XX respectively.
XX (Updated on 25-MAR-2003 to correct PF field.)
XX
XX Sequence 3453 BP; 946 A; 822 C; 876 G; 809 T; 0 other;
XX
Query Match 100.0%; Score 3451.4; DB 16; Length 3453;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3452; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 GCGCCCTGGCTTACCGCGCTCCGAGGCCATGCGGCGTTGGCGCAGCGCAGCGCGG 60
Db 1 GCGCCCTGGCTTACCGCGCTCCGAGGCCATGCGGCGTTGGCGCAGCGCAGCGCGG 60
Qy 61 CGGCTGCTGCTGCTGTTGTTTGTCTAGTAAATGATTTCTTGAGACCGTTACAAACCAAGAC 120
Db 61 CGGCTGCTGCTGCTGTTGTTTGTCTAGTAAATGATTTCTTGAGACCGTTACAAACCAAGAC 120
Qy 121 CTGCTGTGATCAAGTGTGTTTAAATCAGTCATGAGAACATGCTCATCAGCGGGAAG 180
Db 121 CTGCTGTGATCAAGTGTGTTTAAATCAGTCATGAGAACATGCTCATCAGCGGGAAG 180
Qy 181 CCATCATCTGATCCGATGTTGCGAGGATCCCAAGAACCTCCAGTGTATCCCGAGGCGC 240
Db 181 CCATCATCTGATCCGATGTTGCGAGGATCCCAAGAACCTCCAGTGTATCCCGAGGCGC 240
Qy 241 CAGAGTGAAGGACGGTATATGAGCGGCCACCGTGGAGTGGCGGAGTCTGGGTCCATC 300
Db 241 CAGAGTGAAGGACGGTATATGAGCGGCCACCGTGGAGTGGCGGAGTCTGGGTCCATC 300
Qy 301 ACCCTGCAAGTGCAGCTCGCCACCCAGCGGACCTTTCTGCTCTGGGTCTTTAAGCAC 360
Db 301 ACCCTGCAAGTGCAGCTCGCCACCCAGCGGACCTTTCTGCTCTGGGTCTTTAAGCAC 360
Qy 361 AGCTCCCTGGGTGCGAGCGCCACTTTTGAATTTACAAAACAGAGGAATCGTTTCCATGGCC 420
Db 361 AGCTCCCTGGGTGCGAGCGCCACTTTTGAATTTACAAAACAGAGGAATCGTTTCCATGGCC 420

```

QY 421 ATCTTGAACGTACACAGACACAGGACGAGAGATACCTACTCCATATTCAGACGGAACGC 480
DB |||||
QY 421 ATCTTGAACGTACACAGACACAGGACGAGAGATACCTACTCCATATTCAGACGGAACGC 480
DB |||||
QY 481 GCCAACTACACAGTACTGTTTCACAGTGAATGTAAGAGATACACAGCTGTATGTGCTAAGG 540
DB |||||
QY 481 GCCAACTACACAGTACTGTTTCACAGTGAATGTAAGAGATACACAGCTGTATGTGCTAAGG 540
DB |||||
QY 541 AGACCTTACTTTAGGAAGATGAAACCAAGAGTGAATGTAAGAGATACACAGCTGTATGTGCTAAGG 600
DB |||||
QY 541 AGACCTTACTTTAGGAAGATGAAACCAAGAGTGAATGTAAGAGATACACAGCTGTATGTGCTAAGG 600
DB |||||
QY 601 CCGGAGCCACTGTGGAGTGGGTGCTCTCGAGCTCCACAGGGAAGCTGTAAGAGAA 660
DB |||||
QY 601 CCGGAGCCACTGTGGAGTGGGTGCTCTCGAGCTCCACAGGGAAGCTGTAAGAGAA 660
DB |||||
QY 661 GGCCTCTGCTGTGTTCAGAAAGAGGAAAGAGTACTTTCATGAGTGTTCGGAAACAGACATC 720
DB |||||
QY 661 GGCCTCTGCTGTGTTCAGAAAGAGGAAAGAGTACTTTCATGAGTGTTCGGAAACAGACATC 720
DB |||||
QY 721 AGATGCTGTGTAGAAATGCACTGGGCGCGGAATGCAACCAAGCTGTTCACATAGATCTA 780
DB |||||
QY 721 AGATGCTGTGTAGAAATGCACTGGGCGCGGAATGCAACCAAGCTGTTCACATAGATCTA 780
DB |||||
QY 781 AACCAGGCTCTCAGAGCACACTGCCCGAGTTATTCCTGAAAGTGGGGAAACCTTGTGG 840
DB |||||
QY 781 AACCAGGCTCTCAGAGCACACTGCCCGAGTTATTCCTGAAAGTGGGGAAACCTTGTGG 840
DB |||||
QY 841 ATCAGGTGAAGGCCATCATGTGAACCATGGATTCGGGCTCACCTGGGAGCTGGAAGAC 900
DB |||||
QY 841 ATCAGGTGAAGGCCATCATGTGAACCATGGATTCGGGCTCACCTGGGAGCTGGAAGAC 900
DB |||||
QY 901 AAGCCCTCGAGGAGGCGAGCTACTTTGAGATGAGTACCTTCTCCACAAACAGGACCATG 960
DB |||||
QY 901 AAGCCCTCGAGGAGGCGAGCTACTTTGAGATGAGTACCTTCTCCACAAACAGGACCATG 960
DB |||||
QY 961 ATTCCGATTTCTCTTGCCCTTTGTGCTTCCTCGTGGGAAGAAACGACACCGGATATACAC 1020
DB |||||
QY 961 ATTCCGATTTCTCTTGCCCTTTGTGCTTCCTCGTGGGAAGAAACGACACCGGATATACAC 1020
DB |||||
QY 1021 TGCTCTTCTCTAAGCACCCAGCCAGTCAGCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1080
DB |||||
QY 1021 TGCTCTTCTCTAAGCACCCAGCCAGTCAGCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1080
DB |||||
QY 1081 ATAAACGCTTACAGCTTCGCAAGAGAGTATGAAATTCAGCCGTCAGAAAGTTCTGCTTC 1140
DB |||||
QY 1081 ATAAACGCTTACAGCTTCGCAAGAGAGTATGAAATTCAGCCGTCAGAAAGTTCTGCTTC 1140
DB |||||
QY 1141 TCAGTCAGGTTTAAAGCGTACCCACGAATCCGATGCAAGTGGATCTTCTCAAGCCTCA 1200
DB |||||
QY 1141 TCAGTCAGGTTTAAAGCGTACCCACGAATCCGATGCAAGTGGATCTTCTCAAGCCTCA 1200
DB |||||
QY 1201 TTTCTCTGTGAACAGAGAGGCTGGAGGATGGGTACAGCATATCTAAATTTTGGGATCAT 1260
DB |||||
QY 1201 TTTCTCTGTGAACAGAGAGGCTGGAGGATGGGTACAGCATATCTAAATTTTGGGATCAT 1260
DB |||||
QY 1261 AAGAACAGCCAGGAGAGTACATATTTCTATGCAAGAAATGATGACGCCAGTTTCACCAA 1320
DB |||||
QY 1261 AAGAACAGCCAGGAGAGTACATATTTCTATGCAAGAAATGATGACGCCAGTTTCACCAA 1320
DB |||||
QY 1321 ATGTTTCAAGCTGAATATAAGAAAGAAACCTCAAGTGTCTAGCAATGCTCAGCCAGCCAG 1380
DB |||||
QY 1321 ATGTTTCAAGCTGAATATAAGAAAGAAACCTCAAGTGTCTAGCAATGCTCAGCCAGCCAG 1380
DB |||||
QY 1381 GCGTCTGTTCTCTGATGGCTACCCGCTACCTCTTGGACCTGGAAGAGTGTTCGGAC 1440
DB |||||
QY 1381 GCGTCTGTTCTCTGATGGCTACCCGCTACCTCTTGGACCTGGAAGAGTGTTCGGAC 1440
DB |||||
QY 1441 AAATCTCCAAATGACGAGGAAATCCCAAGAGGAGTTTGGAAATATAAGAGGCTAACAGA 1500
DB |||||
QY 1441 AAATCTCCAAATGACGAGGAAATCCCAAGAGGAGTTTGGAAATATAAGAGGCTAACAGA 1500
DB |||||

QY 1501 AAAGTGTTTGGCAGTGGGTGTCGAGCAGTACTCTAAATATAGTGGAGCCGCGGAAGGG 1560
DB |||||
QY 1501 AAAGTGTTTGGCAGTGGGTGTCGAGCAGTACTCTAAATATAGTGGAGCCGCGGAAGGG 1560
DB |||||
QY 1561 CTTCTGGTCAAAATGCTGTCGCTACAAATTCATATGGGCACTCTTTGCGAAACCATCTTTTAA 1620
DB |||||
QY 1561 CTTCTGGTCAAAATGCTGTCGCTACAAATTCATATGGGCACTCTTTGCGAAACCATCTTTTAA 1620
DB |||||
QY 1621 AACTCACAGGCCCCCTTCCCTTTTCATCCAAAGACACATCTCCTTCTATGGACCATTTGG 1680
DB |||||
QY 1621 AACTCACAGGCCCCCTTCCCTTTTCATCCAAAGACACATCTCCTTCTATGGACCATTTGG 1680
DB |||||
QY 1681 CTTCTGCTCCCTCTCAATGCTGCTCTCATTTGTTGATCTGCGACAAATACAAAAAGCAA 1740
DB |||||
QY 1681 CTTCTGCTCCCTCTCAATGCTGCTCTCATTTGTTGATCTGCGACAAATACAAAAAGCAA 1740
DB |||||
QY 1741 TTTAGTACAGAGTCAAGTCAAGTATGATCCAGGTGATCCAGGTGACCTGGCCCTTGGATACGAGTAC 1800
DB |||||
QY 1741 TTTAGTACAGAGTCAAGTCAAGTATGATCCAGGTGATCCAGGTGACCTGGCCCTTGGATACGAGTAC 1800
DB |||||
QY 1801 TTTAGTACAGTCAAGTCAAGTATGATCCAGGTGATCCAGGTGAGTTCGGAGAGAGAAC 1860
DB |||||
QY 1801 TTTAGTACAGTCAAGTCAAGTATGATCCAGGTGATCCAGGTGAGTTCGGAGAGAGAAC 1860
DB |||||
QY 1861 TTTAGAGTGTGGGAAGTCTGCGGCTCTGCGGAGGTGATGAACGCCACCGCC 1920
DB |||||
QY 1861 TTTAGAGTGTGGGAAGTCTGCGGCTCTGCGGAGGTGATGAACGCCACCGCC 1920
DB |||||
QY 1921 TATGGCATTTAGTAAACCGGAGTCTCAATTCAGGTGGCGGTGAAGATGCTTAAAGAGAAA 1980
DB |||||
QY 1921 TATGGCATTTAGTAAACCGGAGTCTCAATTCAGGTGGCGGTGAAGATGCTTAAAGAGAAA 1980
DB |||||
QY 1981 GCTGACAGTGTGAAAGAAAGAGTCTCATGTGCGAGCTCAAAATGATGACCCCTGGGA 2040
DB |||||
QY 1981 GCTGACAGTGTGAAAGAAAGAGTCTCATGTGCGAGCTCAAAATGATGACCCCTGGGA 2040
DB |||||
QY 2041 CACCATGACAAATCGTGAATCTGCTGGGGGATGACACTGTGAGGGGCGAGTGTACTTG 2100
DB |||||
QY 2041 CACCATGACAAATCGTGAATCTGCTGGGGGATGACACTGTGAGGGGCGAGTGTACTTG 2100
DB |||||
QY 2101 ATTTTGAATATTTGCTGATGCTGACCTCTCACTACCTAAGAAAGTAAAGAGAGAA 2160
DB |||||
QY 2101 ATTTTGAATATTTGCTGATGCTGACCTCTCACTACCTAAGAAAGTAAAGAGAGAA 2160
DB |||||
QY 2161 TTTCAAGACATGAGACAGAGATTTTAAAGAACATAATTTTCAAGTCTTACCCTACTTTC 2220
DB |||||
QY 2161 TTTCAAGACATGAGACAGAGATTTTAAAGAACATAATTTTCAAGTCTTACCCTACTTTC 2220
DB |||||
QY 2221 CAGGACATTTCAATTTCCAGCATGCTGTTCAAGAGAGTTCAGTTACACCCGCGCTTG 2280
DB |||||
QY 2221 CAGGACATTTCAATTTCCAGCATGCTGTTCAAGAGAGTTCAGTTACACCCGCGCTTG 2280
DB |||||
QY 2281 GATCAGCTCTCAGGTTCAATGCGGAATTCATTTCTGAGATGAGATGGAATGAA 2340
DB |||||
QY 2281 GATCAGCTCTCAGGTTCAATGCGGAATTCATTTCTGAGATGAGATGGAATGAA 2340
DB |||||
QY 2341 AACCAAGAGAGGCTGGCAGAGAGAGAGAGAGATTTTGAACGCTGCTGAGCTTTGAAGAC 2400
DB |||||
QY 2341 AACCAAGAGAGGCTGGCAGAGAGAGAGAGAGATTTTGAACGCTGCTGAGCTTTGAAGAC 2400
DB |||||
QY 2401 CTTCTTTGCTTTGCTGATCCAAAGTGGCCAAAGGATGGAATTTCTGAGTTCAGTCTGT 2460
DB |||||
QY 2401 CTTCTTTGCTTTGCTGATCCAAAGTGGCCAAAGGATGGAATTTCTGAGTTCAGTCTGT 2460
DB |||||
QY 2461 GTCCACAGAGCTGGCAGCAGGATGTTGGTCCACCCAGGAGGTTGGAAGATC 2520
DB |||||
QY 2461 GTCCACAGAGCTGGCAGCAGGATGTTGGTCCACCCAGGAGGTTGGAAGATC 2520
DB |||||
QY 2521 TGTGACTTTGAGTCTGGCCCGAGACATCTCCAGCTACCTCGTCAAGGGGCAAC 2580
DB |||||
QY 2521 TGTGACTTTGAGTCTGGCCCGAGACATCTCCAGCTACCTCGTCAAGGGGCAAC 2580
DB |||||
QY 2581 GCAAGGCTCCCGTGAAGTGGATGGCAACCGGAGAGCTTATTTGAAGGGATCTACAAATC 2640
DB |||||

Db 2581 GCACGGCTGCGCGTGAAGTGGTGCACCCGAGAGCTTATTTGAAGGATCTACACATC 2640
 Qy 2641 AAGAGTCACGTCCTGCTCCTACGGCATCTCTCTGGGAGATATTTTCACTGGGTGTGAAC 2700
 Db 2641 AAGAGTCACGTCCTGCTCCTACGGCATCTCTCTGGGAGATATTTTCACTGGGTGTGAAC 2700
 Qy 2701 CCTTACCTGGCATCTCTGCTGCGAGCTAACTTCTATAAATGATTTCAGAGTGGATTTAAA 2760
 Db 2701 CCTTACCTGGCATCTCTGCTGCGAGCTAACTTCTATAAATGATTTCAGAGTGGATTTAAA 2760
 Qy 2761 ATGAGCAGCATCTATGCGACAGAGGATATATCTTTGTAATGCAATCTCTGCTGGCT 2820
 Db 2761 ATGAGCAGCATCTATGCGACAGAGGATATATCTTTGTAATGCAATCTCTGCTGGCT 2820
 Qy 2821 TTTGACTCAAGGAGCGGCTCTCTCCCAACCTGACTTCTATTTTAGATGTGAGCTG 2880
 Db 2821 TTTGACTCAAGGAGCGGCTCTCTCCCAACCTGACTTCTATTTTAGATGTGAGCTG 2880
 Qy 2881 GCAGAGCGAGAGCATGTATCAGAAATCCATCCATCTACCAAAACAGGCGGCCCT 2940
 Db 2881 GCAGAGCGAGAGCATGTATCAGAAATCCATCCATCTACCAAAACAGGCGGCCCT 2940
 Qy 2941 CAGCAGAGAGCGGCTCAGAGCCAGTGCACAGCCAGGTGAAGATTACAGAGAA 3000
 Db 2941 CAGCAGAGAGCGGCTCAGAGCCAGTGCACAGCCAGGTGAAGATTACAGAGAA 3000
 Qy 3001 AGAAGTTAGGAGGAGGCTTGACCCCGCCAGCTAGCGGTGAGAGCCGAGAGCCA 3060
 Db 3001 AGAAGTTAGGAGGAGGCTTGACCCCGCCAGCTAGCGGTGAGAGCCGAGAGCCA 3060
 Qy 3061 AGATTAGCTCGCTCTGAGGAGCGCCCTACAGCGGCTGCTTCTGCTGACTTTTCTCT 3120
 Db 3061 AGATTAGCTCGCTCTGAGGAGCGCCCTACAGCGGCTGCTTCTGCTGACTTTTCTCT 3120
 Qy 3121 AGATGCTGTCCATTACTCCAAAGTGACTTCTATAAATCAAACTCTCTCGCACAG 3180
 Db 3121 AGATGCTGTCCATTACTCCAAAGTGACTTCTATAAATCAAACTCTCTCGCACAG 3180
 Qy 3181 GGGGAGAGCCAAATAGAGACTTGTGTGAGCCGCTACCTGCGGGGCTTTCCAG 3240
 Db 3181 GGGGAGAGCCAAATAGAGACTTGTGTGAGCCGCTACCTGCGGGGCTTTCCAG 3240
 Qy 3241 AGCTTGAGGGGAAAGCCATGTATCTGAATATAGTATATTTCTGTAATACTGGAACNA 3300
 Db 3241 AGCTTGAGGGGAAAGCCATGTATCTGAATATAGTATATTTCTGTAATACTGGAACNA 3300
 Qy 3301 ACCAAACCCGTTTTTGTCTAAGGGAAGCTAAATATGATTTTAAATCTATGTTTTAA 3360
 Db 3301 ACCAAACCCGTTTTTGTCTAAGGGAAGCTAAATATGATTTTAAATCTATGTTTTAA 3360
 Qy 3361 AATACTATGTAACCTTTTTCATCTATTAGTGTATATTTTATGATGAAATAAATTTTC 3420
 Db 3361 AATACTATGTAACCTTTTTCATCTATTAGTGTATATTTTATGATGAAATAAATTTTC 3420
 Qy 3421 TACTGTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 3453
 Db 3421 TACTGTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 3453

RESULT 9

AAQ29954
 ID AAQ29954 standard; cDNA; 3453 BP.
 XX
 AC AAQ29954;
 XX
 DT 25-MAR-2003 (updated)
 DT 15-MAR-1993 (first entry)
 XX
 DE Murine flk-2 cDNA sequence.
 XX
 KW Thymidine kinase; TK; haematopoietic; stem cells; proliferation;
 KW differentiation; progenitor cells; foetal liver kinase; ss.

XX OS Mus musculus.
 XX Key Location/Qualifiers
 FT CDS 31..3009
 FT /*tag= a
 XX
 DN W09217486-A1.
 XX
 PD 15-OCT-1992.
 XX
 PF 02-APR-1992; 92WO-US02750.
 XX
 PR 02-APR-1991; 91US-0679666.
 PR 28-JUN-1991; 91US-0728913.
 PR 15-NOV-1991; 91US-0793065.
 PR 24-DEC-1991; 91US-0813593.
 XX
 PA (UYPR-) UNIV PRINCETON.
 XX
 PI Lemischka IR;
 XX
 WI P1; 1992-366185/44.
 DR P-PSDB; AAR28038.
 XX
 PT Stimulating proliferation and/or differentiation of primitive
 PT mammalian haematopoietic stem cells - using ligand that binds
 XX thymidine kinase and flk-1 and flk-2
 PS Claim 6; Fig 1a; 94pp; English.
 XX
 CC The murine foetal liver kinase (flk) -2 clone was isolated by
 CC standard PCR techniques from stem-cell receptor-contg. tissue cDNA
 CC libraries. Suitable tissues include foetal liver, spleen or thymus
 CC cells or adult marrow or brain cells. The PCR primers used are based
 CC on known sections of the flk-2 gene. The murine flk-2 clone may be
 CC used in a vector to transform haematopoietic cells. The thymidine
 CC kinase encoded by flk-2 is expressed in primitive but not mature
 CC haematopoietic cells. Ligand binding to the TK may be prep'd. which
 CC can stimulate proliferation and/or differentiation of primitive
 CC haematopoietic cells in vivo. The ligands can stimulate the
 CC proliferation of additional primitive stem cells, differentiation into
 CC more mature progenitor cells, or both.
 CC See also AAQ29955-7.
 CC (Updated on 25-MAR-2003 to correct PN field.)
 XX
 SQ Sequence 3453 BP; 946 A; 822 C; 876 G; 809 T; 0 other;
 Query Match 99.9%; Score 3448.2; DB 13; Length 3453;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 3450; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 Qy 1 GCGGCTGGCTACCGCGCTCCGAGGCCATCGGGCTTGGCGCAGCGACCGCG 60
 Db 1 GCGGCTGGCTACCGCGCTCCGAGGCCATCGGGCTTGGCGCAGCGACCGCG 60
 Qy 61 CGGCTGCTGCTGCTTGTGTTTGTCTAGTATGTTCTTGAGACCTTACAAACCAAGAC 120
 Db 61 CGGCTGCTGCTGCTTGTGTTTGTCTAGTATGTTCTTGAGACCTTACAAACCAAGAC 120
 Qy 121 CTGCTGTGATCAAGTGTGTTTAAATCAGTATGAGAACCAATGCTCATCAGCGGAAAG 180
 Db 121 CTGCTGTGATCAAGTGTGTTTAAATCAGTATGAGAACCAATGCTCATCAGCGGAAAG 180
 Qy 181 CCATCATCGTACCGAATGGTGGAGGATCCCGAAGACCTCCAGTGTACCCCGAGGCGC 240
 Db 181 CCATCATCGTACCGAATGGTGGAGGATCCCGAAGACCTCCAGTGTACCCCGAGGCGC 240
 Qy 241 CAGAGTGAAGGAGCGGTATATGAGCGGCCACCGTGGAGGTGGCGGATCTGGGTCCATC 300
 Db 241 CAGAGTGAAGGAGCGGTATATGAGCGGCCACCGTGGAGGTGGCGGATCTGGGTCCATC 300
 Qy 301 ACCCTGCAAGTGCAGTCTCGCCACCCAGGGGACCTTTCTCTGCTCTTAAAGCAC 360

[illegible]

Db 2281 GATCAGCTCTCAGGGTTCAATGGGAATTCATTCTCGAAGATGAGATTGAATATGAA 2340
QY 2341 AACCAAGAAGGCTGGCAGAGAAGAGGAGAGATTTGAACCTGCTGAGCTTTGAAGAC 2400
Db 2341 AACCAAGAAGGCTGGCAGAGAAGAGGAGAGATTTGAACCTGCTGAGCTTTGAAGAC 2400
QY 2401 CTCCTTTGCTTTGCTGATACCAAGTGGCCAAAGGATGGAATTCCTGGAGTTCAAGTCTGT 2460
Db 2401 CTCCTTTGCTTTGCTGATACCAAGTGGCCAAAGGATGGAATTCCTGGAGTTCAAGTCTGT 2460
QY 2461 GTCCACAGAGACTGGCAGCAGGAGATGTTGGTCAACCCAGGAGGTTGAAGATC 2520
Db 2461 GTCCACAGAGACTGGCAGCAGGAGATGTTGGTCAACCCAGGAGGTTGAAGATC 2520
QY 2521 TGTGACTTTGGACTGGCCCGCAGACATCTCGAGGACTCCAGCTACCTGTCAGGGCAAC 2580
Db 2521 TGTGACTTTGGACTGGCCCGCAGACATCTCGAGGACTCCAGCTACCTGTCAGGGCAAC 2580
QY 2581 GCACGGCTGCCGCTGAAGTGGATGTCACCCAGGAGCTTATTTGAAGGATCTACACAATC 2640
Db 2581 GCACGGCTGCCGCTGAAGTGGATGTCACCCAGGAGCTTATTTGAAGGATCTACACAATC 2640
QY 2641 AAGAGTACCTGCTGCTACCGCATCTCTCTCGGAGATATTTTCACTGGGTGTGAAC 2700
Db 2641 AAGAGTACCTGCTGCTACCGCATCTCTCTCGGAGATATTTTCACTGGGTGTGAAC 2700
QY 2701 CCTTACCTTGGCATCTCTGTCAGCCTAACTTCTATAACTGATTCAGAGTGGATTTAA 2760
Db 2701 CCTTACCTTGGCATCTCTGTCAGCCTAACTTCTATAACTGATTCAGAGTGGATTTAA 2760
QY 2761 ATGAGAGCCTATCTATGCCACAGAGGATATCTTTGTAATGCAATCCCTGCTGGCT 2820
Db 2761 ATGAGAGCCTATCTATGCCACAGAGGATATCTTTGTAATGCAATCCCTGCTGGCT 2820
QY 2821 TTTGACTCAAGAGCGGCTCTTCCCAACCTGACTTCAATTTTAGGATGTGAGCTG 2880
Db 2821 TTTGACTCAAGAGCGGCTCTTCCCAACCTGACTTCAATTTTAGGATGTGAGCTG 2880
QY 2881 GCAGAGGAGAGAGCATGATCAGACATCCATCCATCCATCCATCCATCCATCCATCCAT 2940
Db 2881 GCAGAGGAGAGAGCATGATCAGACATCCATCCATCCATCCATCCATCCATCCATCCAT 2940
QY 2941 CAGCAGAGAGCGGCTCAGAGCCAGTCCGACAGCCAGGTCAGAGTTCAGAGGACCA 3000
Db 2941 CAGCAGAGAGCGGCTCAGAGCCAGTCCGACAGCCAGGTCAGAGTTCAGAGGACCA 3000
QY 3001 AGAAGTTAGCAGAGGCTTTGGACCCCGCCACCTAGCAGGCTGTAGACCCGAGAGCCA 3060
Db 3001 AGAAGTTAGCAGAGGCTTTGGACCCCGCCACCTAGCAGGCTGTAGACCCGAGAGCCA 3060
QY 3061 AGATTAGCCTCGCTCTGAGAGAGGCGCTTACAGGCGTTCGCTGCTGAGCTTTCTCT 3120
Db 3061 AGATTAGCCTCGCTCTGAGAGAGGCGCTTACAGGCGTTCGCTGCTGAGCTTTCTCT 3120
QY 3121 AGATGCTGTCTGCCATTACTCCTAAAGTGAATTTCTATAAATCAAACTCTCTCGCAG 3180
Db 3121 AGATGCTGTCTGCCATTACTCCTAAAGTGAATTTCTATAAATCAAACTCTCTCGCAG 3180
QY 3181 GCGGGAGAGCAATATAGACTTTGTTGGTGGAGCCCGCTACCTGCGGGCTTTTCCACG 3240
Db 3181 GCGGGAGAGCAATATAGACTTTGTTGGTGGAGCCCGCTACCTGCGGGCTTTTCCACG 3240
QY 3241 AGCTTGAGGGGAAGCCATGATCTGAATATAGTATATCTTGTAAATACGTGAACAA 3300
Db 3241 AGCTTGAGGGGAAGCCATGATCTGAATATAGTATATCTTGTAAATACGTGAACAA 3300
QY 3301 ACCAAACCCCTTTTGTCTAAGGAAAGCTAAATATGATTTTAAATCTATGTTTTAA 3360
Db 3301 CCAAAACCCCTTTTGTCTAAGGAAAGCTAAATATGATTTTAAATCTATGTTTTAA 3360
QY 3361 AATATGATGTAACCTTTTTCATCTATTTAGTGATATTTTATGATGGAATTAACCTTC 3420
Db 3361 AATATGATGTAACCTTTTTCATCTATTTAGTGATATTTTATGATGGAATTAACCTTC 3420

QY 3421 TACTGTAAAAA 3453
Db 3421 TACTGTAAAAA 3453

RESULT 11

AAD42483
ID AAD42483 standard; cDNA; 3440 BP.

XX AAD42483;

DT 15-NOV-2002 (first entry)

XX Murine receptor protein tyrosine kinase, flk-2 cDNA.

XX Murine; receptor protein tyrosine kinase; pTK; haematopoietic cell;
KW growth factor; therapeutic; macrocytic anaemia; aplastic anaemia;
KW bone marrow damage; cancer; chemotherapy; radiation; osteopathic;
KW flk-2; gene; ss.

XX Mus sp.

XX Key Location/Qualifiers
FT CDS 31..3009

FT /tag= a
FT /product= "Murine flk-2 protein"

FT sig_peptide 31..111

FT mat_peptide 112..3006

FT /tag= c

FT /product= "Murine mature flk-2 protein"

FT /tag= d

FT /note= "Encodes extracellular domain"

FT /tag= e

FT /note= "Encodes transmembrane domain"

FT /tag= f

FT /note= "Encodes intracellular domain"

XX US2002072077-A1.

XX 13-JUN-2002.

XX 31-JUL-2001; 2001US-0919408.

XX 19-NOV-1992; 92US-0977451.

XX 30-APR-1993; 93US-0055269.

XX 31-OCT-1994; 94US-0252498.

XX 15-FEB-1996; 96US-0601891.

XX 10-FEB-1998; 98US-0021324.

XX 10-DEC-1998; 98US-0208786.

XX 02-APR-1991; 91US-0679666.

XX 28-JUN-1991; 91US-0728913.

XX 15-NOV-1991; 91US-0793065.

XX 24-DEC-1991; 91US-0813593.

XX (LEMI/) LEMISCHKA I R.

XX Lemischka IR;

XX WPI; 2002-607237/65.

XX P-PSDB; AAE25818.

XX New protein tyrosine kinase expressed in primitive hematopoietic cells
(HC) and not expressed in mature HC, and ligands for the protein, for
stimulating proliferation of primitive hematopoietic stem cells
Claim 6; Page 15-19; 64pp; English.
The present invention relates to receptor protein tyrosine kinase (pTK)

1801	DB	TTCTACGTTGACTTCAGGACATATGAATATGACCTTAAGTGGAGTTC	1860	GGAGAGAGAAC
1861	QY	TTAGAGTTTGGGAAGTCCCTGGGGTCTGGCGCTTTCCGGAGGGTGATGAA	1920	CGCCACGCCC
1861	DB	TTAGAGTTTGGGAAGTCCCTGGGGTCTGGCGCTTTCCGGAGGGTGATGAA	1920	CGCCACGCCC
1921	QY	TATGGCATTATGATAAAACGGGAGTCTCAATTCAGTGGCGGTGAAGATGCT	1980	TAAGAGAGAAA
1921	DB	TATGGCATTATGATAAAACGGGAGTCTCAATTCAGTGGCGGTGAAGATGCT	1980	TAAGAGAGAAA
1981	QY	GCTGACAGCTGTGAAAAAGAGAGCTCTCATGTCCGAGCTCAAAATGATGAC	2040	CCACCTGGGA
1981	DB	GCTGACAGCTGTGAAAAAGAGAGCTCTCATGTCCGAGCTCAAAATGATGAC	2040	CCACCTGGGA
2041	QY	CACCATGACAAATCGTGAATCTGCTGGGGGATGCAACATGTCAGGGCCAGT	2100	GCTACTTG
2041	DB	CACCATGACAAATCGTGAATCTGCTGGGGGATGCAACATGTCAGGGCCAGT	2100	GCTACTTG
2101	QY	ATTTTTGAAATATGTTGCTATGTTGACCTCTCTCAACTTAAGAGATGAAAG	2160	AGAGAGAGG
2101	DB	ATTTTTGAAATATGTTGCTATGTTGACCTCTCTCAACTTAAGAGATGAAAG	2160	AGAGAGAGG
2161	QY	TTTCACAGGACATGACAGAGATTTTTTAAGAAATATATTCAGTCTTACCCT	2220	CTACTTTC
2161	DB	TTTCACAGGACATGACAGAGATTTTTTAAGAAATATATTCAGTCTTACCCT	2220	CTACTTTC
2221	QY	CAGGCACATTCAAATTCAGAGATCGCTGGTTTACGAGAAAGTTACACCCG	2280	CGCCCTTG
2221	DB	CAGGCACATTCAAATTCAGAGATCGCTGGTTTACGAGAAAGTTACACCCG	2280	CGCCCTTG
2281	QY	GATCAGCTCTCAGGGTTCAATGAGAAATTCATTCAGATGAGATGGAATGAA	2340	TATGAA
2281	DB	GATCAGCTCTCAGGGTTCAATGAGAAATTCATTCAGATGAGATGGAATGAA	2340	TATGAA
2341	QY	AACACAGAGAGCTGGCAGAGAGAGAGAGAGATTTGAACGTCTGACGTTGA	2400	AGAGAC
2341	DB	AACACAGAGAGCTGGCAGAGAGAGAGAGAGATTTGAACGTCTGACGTTGA	2400	AGAGAC
2401	QY	CTCCTTTGCTTTGCGTATCAAGTGGCCAAAGGCATGAAATTCCTGGAGTT	2460	CAAGTCGTG
2401	DB	CTCCTTTGCTTTGCGTATCAAGTGGCCAAAGGCATGAAATTCCTGGAGTT	2460	CAAGTCGTG
2461	QY	GTCCACAGAGACCTGGCAGCCAGGAATGTGTGTGTCCACCGGAGGTGTGA	2520	AGATC
2461	DB	GTCCACAGAGACCTGGCAGCCAGGAATGTGTGTGTCCACCGGAGGTGTGA	2520	AGATC
2521	QY	TGTGACTTTGCACTGGCCCGAGACATCCTCAGCGACTCAGCTACGTCTG	2580	CGGGGCAAC
2521	DB	TGTGACTTTGCACTGGCCCGAGACATCCTCAGCGACTCAGCTACGTCTG	2580	CGGGGCAAC
2581	QY	GCACGGCTGCGGTGAAGTGGATGGCACCCGAGAGCTTATTTGAAGGATCT	2640	ACACATC
2581	DB	GCACGGCTGCGGTGAAGTGGCACCCGAGAGCTTATTTGAAGGATCTACAC	2640	ATCAATC
2641	QY	AAGAGTGAACCTCTGCTTACGGCATCCCTCTCTGGGAGATATTTTCACT	2700	GGGTGTGAAC
2641	DB	AAGAGTGAACCTCTGCTTACGGCATCCCTCTCTGGGAGATATTTTCACT	2700	GGGTGTGAAC
2701	QY	CTTTACCTTGGCATCTCTGTGCAAGCTTAACTTCTATAAACTGATTCAG	2760	AGTGAATTTAAA
2701	DB	CTTTACCTTGGCATCTCTGTGCAAGCTTAACTTCTATAAACTGATTCAG	2760	AGTGAATTTAAA
2761	QY	ATGAGCAGGCATTTCTATGCCACAGAGGATATATCTTTGTAATGCAATC	2820	CTGCTGGCT
2761	DB	ATGAGCAGGCATTTCTATGCCACAGAGGATATATCTTTGTAATGCAATC	2820	CTGCTGGCT
2821	QY	TTTGACTCAAGGAAGCGGCCATCCTTCCGCAACCTGACTTCAATTTT	2880	TAGGATGTACGCTG
2821	DB	TTTGACTCAAGGAAGCGGCCATCCTTCCGCAACCTGACTTCAATTTT	2880	TAGGATGTACGCTG
2881	QY	GCAGGGCAGAGAACATGATACAGAACTCATCATCTACCAAAACAGGGC	2940	CGCCCT
2881	DB	GCAGGGCAGAGAACATGATACAGAACTCATCATCTACCAAAACAGGGC	2940	CGCCCT

RESULT 12

AAT00801

AA100001
ID AAT00801 standard; DNA: 3521 BP.

AA
AC
AAT00801;

DT 29-FEB-1996 (first entry)

DE Flk2/flt3 tyrosine kinase receptor gene.

KW Tyrosine kinase receptor: flk2: flk3: agonist: monoclonal antibody:

haematopoiesis; hypoplasemia; anaemia; thrombocytopenia; stem cell: ss

OS Mus sp.

XX	Key	Location/Qualifiers
FH		

FT CDS 82..3084

```
FT
/*tag= a
```

PN WO9527062-A1.

PD 12-OCT-1995.

PF 23-MAR-1995; 95WO-US03718.

PR 04-APR-1994; 94US-02222299.

PA (GETH) GENENTECH INC.

XX
PI
Bennett BD. Broz SD Matthews W Zolotar EC.

XX
DP WPT: 1005-3E9C3C/AC

DR WPI; 1995-358636/
DR P-PSDB: AAB91868

XX Agonist antibody which binds to flk2-flt3 tyrosine kinase receptor

QY	1921	TATGGCATTAGTAAACGGGAGTCTCAATTCAAGTGGCGGTGAAGATGCTTAAAGAGAAA	1980
Db	1972	TATGGCATTAGTAAACGGGAGTCTCAATTCAAGTGGCGGTGAAGATGCTTAAAGAGAAA	2031
QY	1981	GCTGACAGCTGTGAAAAAGAAAGCTCTCATGTCCGAGCTCAAAAATGATGACCCACTCGGGA	2040
Db	2032	GCTGACAGCTGTGAAAAAGAAAGCTCTCATGTCCGAGCTCAAAAATGATGACCCACTCGGGA	2091
QY	2041	CACCATGACAAACATCGTGAATCTGCTGGGGGATGACACATGTGAGGGCCAGTGTACTTGT	2100
Db	2092	CACCATGACAAACATCGTGAATCTGCTGGGGGATGACACATGTGAGGGCCAGTGTACTTGT	2151
QY	2101	ATTTTTTGAATATGTGTCTATGGTGACCTCTCAACTACCTTAAAGAGTAAAGAGAGAAG	2160
Db	2152	ATTTTTTGAATATGTGTCTATGGTGACCTCTCAACTACCTTAAAGAGTAAAGAGAGAAG	2211
QY	2161	TTTTCACAGGACATGACACAGAGATTTTTTAAAGAAACATATTTTCAGTCTTATACCTCACTTTC	2220
Db	2212	TTTTCACAGGACATGACACAGAGATTTTTTAAAGAAACATATTTTCAGTCTTATACCTCACTTTC	2271
QY	2221	CAGGCACATTCAAATTCACAGCATGCTGGTTTACGAGAAAGTTACAGTTCACACCGGCCCTTG	2280
Db	2272	CAGGCACATTCAAATTCACAGCATGCTGGTTTACGAGAAAGTTACAGTTCACACCGGCCCTTG	2331
QY	2281	GATCAGCTCTCAGGGTTTCAATGGGAATTCAAATTCATTCTGAAGATGAGATGAATATGAA	2340
Db	2332	GATCAGCTCTCAGGGTTTCAATGGGAATTCAAATTCATTCTGAAGATGAGATGAATATGAA	2391
QY	2341	AACCAGAAAGAGCTGGCAGAAAGAGAGAGGAGAGATTTTGAACGTCTGACGTTTGAAGAC	2400
Db	2392	AACCAGAAAGAGCTGGCAGAAAGAGAGAGGAGAGATTTTGAACGTCTGACGTTTGAAGAC	2451
QY	2401	CTCCTTTTGGCTTACCAAGTGGCCAAAGGCAATGGAATTCCTGGAGTTCAAAGTCTGTGT	2460
Db	2452	CTCCTTTTGGCTTACCAAGTGGCCAAAGGCAATGGAATTCCTGGAGTTCAAAGTCTGTGT	2511
QY	2461	GTCCACAGAGACCTGGCAGCCAGCAAGATGTGTGTGTCAACCCAGGGAAGTGTGTGAAGATC	2520
Db	2512	GTCCACAGAGACCTGGCAGCCAGCAAGATGTGTGTGTCAACCCAGGGAAGTGTGTGAAGATC	2571
QY	2521	TGTGACTTTTGACTGGCCCGAGACATCCTCAGCGGACTTCCAGTACGTCTGTCAGGGGCAAC	2580
Db	2572	TGTGACTTTTGACTGGCCCGAGACATCCTCAGCGGACTTCCAGTACGTCTGTCAGGGGCAAC	2631
QY	2581	GCACGGCTGCGGTTGAAGTGGATGGCACCGGAGAGCTTATTTGAAGGGATCTACACAAATC	2640
Db	2632	GCACGGCTGCGGTTGAAGTGGATGGCACCGGAGAGCTTATTTGAAGGGATCTACACAAATC	2691
QY	2641	AAGAGTACAGTCTGGTCTTAGGCAATCCTTCTCTGGGAGATATTTTCACTGGGTGTGAAC	2700
Db	2692	AAGAGTACAGTCTGGTCTTAGGCAATCCTTCTCTGGGAGATATTTTCACTGGGTGTGAAC	2751
QY	2701	CCTTTACCCTGGCATTTCTGTGCGAGCTTAACTTATATAAAGTATTCAGAGTGGATTTAAA	2760
Db	2752	CCTTTACCCTGGCATTTCTGTGCGAGCTTAACTTATATAAAGTATTCAGAGTGGATTTAAA	2811
QY	2761	ATGGAGCAGCCATTTCTATGCCACAGAGGGATATCTTTGTAATGCAATCTCTGCTGGGCT	2820
Db	2812	ATGGAGCAGCCATTTCTATGCCACAGAGGGATATCTTTGTAATGCAATCTCTGCTGGGCT	2871
QY	2821	TTTTGACTCAAGGAAGGGGCCATCCTTCCCAACCTGACTTCAATTTTATAGGATGTCAAGCTG	2880
Db	2872	TTTTGACTCAAGGAAGGGGCCATCCTTCCCAACCTGACTTCAATTTTATAGGATGTCAAGCTG	2931
QY	2881	GCAGAGCGAAGAAAGC-----ATGTATCAGAAATCATCATCC	2917
Db	2932	GCAGAGCGAAGAAAGCGATGTATCAGAAATCATGGTGGCAACGTCCCAAGAAATCATCATCC	2991
QY	2918	ATCTACCAAAACAGCGGGGCCCTTCAGCAGAGAGGGGGCTCAGAGGC-CAGTGCACAG	2976
Db	2992	ATCTACCAAAACAGCGGGGCCCTTCAGCAGAGAGGGGGCTCAGAGGCATCGGCCACAG	3051

QY	2977	GC	CAGG	TGAAGATTC	CAGAGA	AAAGATT	TAGCG	AGAGG	CCCTT	GGAC	CCCGCC	CACCC	3033
Db	3052	GC	CAGG	TGAAGATTC	CAGAGA	AAAGATT	TAGCG	AGAGG	CCCTT	GGAC	CCCGCC	CACCC	3111
QY	3037	AG	CAGG	CTGTAG	ACCGCAG	AGCCCAAG	TATG	CTCG	CTCT	GAGGA	AGCGCC	CTAC	3096
Db	3112	AG	CAGG	CTGTAG	ACCGCAG	AGCCCAAG	TATG	CTCG	CTCT	GAGGA	AGCGCC	CTAC	3171
QY	3097	CG	TGCTT	CGCTGG	ACATTTT	CTCTAG	ATGCTGT	CTGC	ATTACT	TCCAA	AGTGA	CTTCT	3156
Db	3172	CG	TGCTT	CGCTGG	ACATTTT	CTCTAG	ATGCTGT	CTGC	ATTACT	TCCAA	AGTGA	CTTCT	3231
QY	3157	AA	ATCA	AACTCT	CTCTCG	CACAG	CGGGAG	AGCCAA	TATG	AGACT	TGTG	TGTG	3216
Db	3232	AA	ATCA	AACTCT	CTCTCG	CACAG	CGGGAG	AGCCAA	TATG	AGACT	TGTG	TGTG	3291
QY	3217	GC	CTAC	CCCTGG	GGCCCTT	-----	CCAG	AGCTT	GAGG	GGAA	AGCCAT	GTAT	3268
Db	3292	GC	CTAC	CCCTGG	GGCCCTT	-----	CCAG	AGCTT	GAGG	GGAA	AGCCAT	GTAT	3351
QY	3269	AT	ATAG	TATAT	CTCTGT	GAATAC	GTGA	AAAC	CAAC	CCCG	CTTGC	TAA	3328
Db	3352	AT	ATAG	TATAT	CTCTGT	GAATAC	GTGA	AAAC	CAAC	CCCG	CTTGC	TAA	3411
QY	3329	CT	AAAT	TATG	ATTTT	TAAAT	CTAT	GT	TTTAA	AAAT	ACTAT	GTAA	3388
Db	3412	CT	AAAT	TATG	ATTTT	TAAAT	CTAT	GT	TTTAA	AAAT	ACTAT	GTAA	3471
QY	3389	GT	GAT	TATAT	TTTAT	GATG	GAATA	ATAA	CTTCT	CTG	TAAAA	AAAAA	3438
Db	3472	GT	GAT	TATAT	TTTAT	GATG	GAATA	ATAA	CTTCT	CTG	TAAAA	AAAAA	3521

RESULT 13

AAQ81013

ID AAQ81013 standard; cDNA; 3501 BP.

XX AAQ81013;

XX 25-MAR-2003 (updated)

DT 18-AUG-1995 (first entry)

XX

DE Flk2 receptor protein-tyrosine-kinase cDNA.

XX Human Flk2; receptor protein-tyrosine-kinase; primitive

KW hematopoietic cell; fetal liver kinase; ds.

XX Homo sapiens.

OS

XX

Key Location/Qualifiers

FT 58..3039

FT /tag= a

FT 58..138

FT /product= Flk2 receptor protein-tyrosine-kinase

FT sig_peptide

FT 139..3036

FT /tag= b

FT mat_peptide

FT /tag= c

XX

XX WO9500554-A2.

XX

XX

XX 05-JAN-1995.

XX

XX 17-JUN-1994; 94WO-US06944.

XX

XX 18-JUN-1993; 93US-0080244.

PR 21-JUN-1993; 93US-0081508.

PR 23-NOV-1993; 93US-0157490.

XX

XX (UYPR-) UNIV PRINCETON.

PA

XX

XX Lemischka IR;

XX WPI; 1995-052014/07.

DR

PR 15-NOV-1991; 91US-0793065.
 PR 24-DEC-1991; 91US-0813593.
 PR 26-JUN-1992; 92US-0906397.
 PR 12-NOV-1992; 92US-0975049.
 PR 19-NOV-1992; 92US-0977451.
 PR 30-APR-1993; 93US-0055269.
 XX (UYPR-) UNIV PRINCETON.
 PA Lemischka IR;
 XX WPI; 1995-005894/01.
 DR P-PSDB; AAR67536.
 XX Murine flk-2 receptor protein tyrosine kinase - used to stimulate proliferation and/or stimulation of primitive mammalian haematopoietic stem cells in vitro or in vivo.
 PT Disclosure; Fig. 2A-1F; 69pp; English.
 XX cDNAs encoding receptor protein tyrosine-kinases, mouse fetal liver kinase-2 (flk-2), human flk-2 and mouse flk-1 are given in AAQ79068-70, CC respectively, and the deduced amino acid sequences in AAR67535-37, CC respectively.
 CC (Updated on 25-MAR-2003 to correct PF field.)
 XX
 SQ Sequence 3501 BP; 1068 A; 709 C; 784 G; 940 T; 0 other;
 Query Match 64.5%; Score 2227.2; DB 16; Length 3501;
 Best Local Similarity 79.6%; Pred. No. 0;
 Matches 2775; Conservative 0; Mismatches 653; Indels 56; Gaps 10;
 QY 8 GGCTACCGCGCTCCGAGGCGCATGCGGGCTTGGCGCAGCGACGCGCGGCTGC 67
 DB 35 GGGGACCCCGGCTCCGAGGCGCATGCGGGCTTGGC---GGCGAGCGGCGCACCGTGC 91
 QY 68 TGCTGCTTGTGTTTGTGAGTAATGATCTTGAGACCGTTACAAACCGACCTGCGCTG 127
 DB 92 CGCTGCTGCTGTTTGTGCAATGATATTTGGGACTATTACAAATCAAGATCTGCGCTG 151
 QY 128 TGATCAAGTGTGTTTAAATCAGTCATGAGAACAAATGGCTCATCAGCGGGAAGCCATCAT 187
 DB 152 TGATCAAGTGTGTTTAAATCAATCAATGAGAACAAATGATTCATCAGTGGGAAGTCAATCAT 211
 QY 188 CGTACCGAATGTGGAGGATCCCGAGAGACCTCCAGTGTACCCCGAGCGGCCAGAGTG 247
 DB 212 CATATCCCATGGTATCAGATCCCGGAGACCTCGGGTGTGGTGTGAGACCCAGAGCT 271
 QY 248 AAGGACCGGTATGAAAGCGGCACCGTGGAGGTGGCGAGTCTGGGTCCATCAGCCCTGC 307
 DB 272 CAGGACAGTGTACGAAGCTGCGGTGTGGAAGTGGATGTATCTGCTTCCATCAGACTGC 331
 QY 308 AAGTGCAGTCCGACCCAGGGGACCTTCTGCTCTGGGTCTTAAAGCAGAGCTCCC 367
 DB 332 AAGTGCAGTCCGACCCAGGGGACCTTCTGCTCTGGGTCTTAAAGCAGAGCTCCC 391
 QY 368 TGGGTGCGCAGCGGCATTTGATTTTACAAAAAGAGGAAATCGTTTCCATGGCCATCTTGA 427
 DB 392 TGAATGGCAGCCACATTTGATTTTACAAAAAGAGGAGTGTTCATGTTGATTTGA 451
 QY 428 ACGTGACAGAGACCCAGGAGAGAAATACCTACTCATATTCAGAGCGAACCGGCAACT 487
 DB 452 AAATGACAGAAACCCCAAGCTGGAGAAATACCTACTTTTATTCAGAGTGAAGTACCAAT 511
 QY 488 ACACAGTACTGTTACAGTGAATGTAAGAGATACACAGCTGTATGTCTAAGGAGACCTT 547
 DB 512 ACACAATATGTTTACAGTGAATGTAAGAAATACCTGCTTTTACACATTTAAGAGACCTT 571
 QY 548 ACTTTAGGAAGTGAAGAACAGAGTGCATCTGCTCTGCAATCTCGAGGGTGTTCGGGAGC 607
 DB 572 ACTTTAGAAAATGGAAGAACAGAGCGCTGCTGCTGCTGCAATCTGAGAGCGTTCAGAGC 631
 QY 608 CCACTGTGGAGTGGTGTCTGTCAGCTCCACAGGGGAAAGCTGTAAAGAAAGAGGCCCTG 667

DB 632 CGATCGTGAATGGTGTCTTGGGATTACAGGGGGAAGCTGTAAAGAAAGTCCAG 691
 QY 668 CTGTTGTGAGAAAGAGGAAAGGTAATCTTATGAGTGTTCGGAAACAGACATCAGATGCT 727
 DB 692 CTGTTGTTAAAAAGAGGAAAGGTAATCTTATGAGTGTTCGGAAAGTCCAGTAAAGTGT 751
 QY 728 GTGCTAGAAATGCACTGGGCGCGAATGCACCAAGCTGTTCACCATAGATCTAAACAGG 787
 DB 752 GTGCCAGAAATGAACCTGGGAGGGAATGCACAGGCTGTTCACATAGATCTAAATCAAA 811
 QY 788 CTCCTCAGAGCACACTGCCCCAGTTTATCTCAAGTGGGGAAACCCCTTGTGATCAGGT 847
 DB 812 CTCCTCAGACCACTGGGAGGGAATTTTCTTAAAGTAGGGGAACCCCTTATGATTAAGT 871
 QY 848 GTAAGGCCATCCATGTGAACCATGAGTTCGGGCTCACCTGGAGGTGGGAAGCAAAAGCCC 907
 DB 872 GCNAAGCTGTTTATGTGAACCATGAGTTCGGGCTCACCTGGGAATTTAGAAAAAAGAAC 931
 QY 908 TGGAGGAGGCGAGCTACTTTGAGATGAGTACCTACTCCACAAACAGGACCATGATTCGGA 967
 DB 932 TCGAGGAGGCGAATCTACTTTGAGATGAGTACCTATTTCAACAAACAGAACTATGATACG 991
 QY 968 TTCTCTGCGCTTGTGCTTCCGTGGGAAGAACGACACCGCATATTAACACTGCTCTT 1027
 DB 992 TTCTGTTGCTTTGTATCATCAGTGGCAAGAACGACACCGGATCTACTACTTGTCTT 1051
 QY 1028 CTTCAAGCACCCAGCGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1087
 DB 1052 CTTCAAGCATCCAGTCAATCAGCTTTGGTTTACCATCGTAGGAAAGGATTTATAATG 1111
 QY 1088 CTACAGCTCGCAAGAGATGAAATGACCCCGTACGAAAGTTCCTCTCTCAGTCA 1147
 DB 1112 CTACCAATTCAGTGAAGATGAAATGACCAATATGAAGAGTTCCTCTCTCTGCTCA 1171
 QY 1148 GGTAAAGCGTACCCAGCAATCCGATGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1207
 DB 1172 GGTAAAGCGTACCCAGCAATCCGATGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1231
 QY 1208 GTAAACAGAGAGGCTGGAGGATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1267
 DB 1232 GTAGCAAAAGGCTCTTGTATACGATACAGATATCCAGTTCATCATAGCACC 1291
 QY 1268 AGCCAGAGAGTACATATTCATGAGAAATGATGAGCCCGAGTTCACCAAAATGTTCA 1327
 DB 1292 AGCCAGAGAGTATATATTCATGAGAAATGATGATGCCCCCAATTTACCAAAATGTTCA 1351
 QY 1328 CGCTGAATATGAAGAAACCTCAAGTGTAGCAATGCTCAGCCAGCGAGGCTCT 1387
 DB 1352 CGCTGAATATGAAGAAACCTCAAGTGTCTCGCAAGAGCATCGGCAAGTCTAGGCTCT 1411
 QY 1388 GTTCTCTGATGGCTTACCGCTTACCTCTTGGACCTGGAAGAGTGTTCGGCAAAATCTC 1447
 DB 1412 GTTCTCTGATGGTATCCCATTTACCATCTTGGACCTGGAAGAGTGTTCAGACAGTCTC 1471
 QY 1448 CCAATTCAGCGAGGAAATCCAGAGGAGTGTGGAATAAAAGGCTAACAGAAAGTGT 1507
 DB 1472 CCAATTCAGAGAGATCAGAGAGGAGTGTGGAATAAGAGGCTAACAGAAAGTGT 1531
 QY 1508 TTGGCAGTGGTGTGAGAGTACTCTAAATATGAGTGAAGCCCGGGAAGGCTTCTGG 1567
 DB 1532 TTGGCAGTGGTGTGAGAGTACTCTAAATATGAGTGAAGCCATTAAGAGGCTTCTGG 1591
 QY 1568 TCAATGCTGTGCGTACAAATTTCTATGGGACGCTTTGGGAAACCATCTTTTAAACTCAC 1627
 DB 1592 TCAAGTGTGTGATACAAATTTCCCTTGGCAATCTTGTGAGACGATCTTTTAAACTCTC 1651
 QY 1628 CAGGCCCTTCCCTTTCATCCAAAGCAACATCTCTCTTATCGGACCATTTGGGCTCTGTC 1687
 DB 1652 CAGGCCCTTCCCTTTCATCCAAAGCAACATCTCTTATGCAACAATTTGGTGTGTC 1711
 QY 1688 TCCCTTCAATGTTGTTCTCATTTGTTGATCTGCCCAAAATACAAAAAGCAATTTAGT 1747

1448 CCAATTGACGAGGAAATCCAGAGGAGTTTGGAAATAAAAGGCTAAACAGAAAAGTGT 1507
1472 CCAACTGCACAGAGAGATCACAGAGAGAGTCTGGAATAGAAAGCTTAAACAGAAAAGTGT 1531
1508 TTGGCCAGTGGGTGTCGAGCAGTACTCTAAATATGAGTGGCGGGAAGGGCTTCTGG 1567
1532 TTGGACAGTGGGTGTCGAGCAGTACTCTAAACATGAGTGAAGCCATAAAGGGTTCCTGG 1591
1568 TCAAAATGCTGGGTGATCAAAATCTATGGGCGGTCTTGGAAACCAATCTTTTAAACTCAC 1627
1592 TCAAGTGTCTGATACAAATTCCTCTGGGCATCTCTGTGAGCGATCCCTTTTAAACTCTC 1651
1628 CAGGCCCTTCCCTTTTCATCAAGACAAACATCTCTCTATGCGACCAATGGGCTGTGC 1687
1652 CAGGCCCTTCCCTTTTCATCCAGACAAACATCTCAITATGCAACAAITGGGTGTTGTC 1711
1688 TCCCTTCAATGTTGTTCTCAITGTTGATGTCACAAATACAAAGCAAACTTAGGT 1747
1712 TCCCTTCAATGTTGTTCTCAITGTTGATGTCACAAATACAAAGCAAACTTAGGT 1771
1748 ACGAGAGTCAGTGCAGATGATCAGGTGACTGGCCCCCTGGATACAGGATCTTCTACG 1807
1772 ATGAAAGCCAGCTACAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1831
1808 TTGACTTTCAGGACTATGAATATGACTTAAGTGGGAGTTTCCGAGAGAGAACTTAGGT 1867
1832 TTGATTTTCAGAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1891
1868 TTGGGAAGTCTCGGGTCTGGGCTTTTCGGGAGGAGTGAACCGCCAGCCCTATGGCA 1927
1892 TTGGGAAGTCTAGGATCAGGTGCTTTTGGAAAGTGAAGCAACAGCTTATGGAA 1951
1928 TTAGTAAACGGAGTCTCAATTCAGGTGGCGTGAAGATGCTTAAAGAGAAAGCTGACA 1987
1952 TTAGCAAAAACAGAGTCTCAATTCAGGTGGCGTCAAAATGCTGAAAGAGAAAGCAGACA 2011
1988 GCTGTCAAAAAGAGCTCTCATGCTCGAGCTCAAAATGATGACCCACCTGGGACACATG 2047
2012 GCTGTCAAAAAGAGGACCTCATGCTCAGAACTCAAGATGATGACCCAGCTGGGAAGCCAG 2071
2048 ACAACATCGTAATCTGCTGGGGGACATGACACTGTCCAGGCGCAGTGTACTGATTTTGG 2107
2072 AGAATATTGTGAACCTGCTGGGGCGTGACACTGTCCAGGACCAATTTACTGATTTTGG 2131
2108 AATATTGTTGCTATGTTGACCTCTCAACTCTAAGAGAGTAAAGAGAGAAATTTTACA 2167
2132 AATACTGTTGCTATGTTGATCTTCTCAACTCTAAGAGAGTAAAGAGAGAAATTTTACA 2191
2168 GGACATGGACAGAGATTTTAAAGGACATAATTTCACTTCTTACCCCTACTTTCCAGGAC 2227
2192 GGACTTGGACAGAGATTTTCAAGGAAACAAATTTCACTTCTTACCCCTACTTTCCCAATCAC 2251
2228 APTCAAATTCAGCATGCTGTTTCAAGAGAGTTCAGTTTACACCCGCGCTTGGATCAGC 2287
2252 ATCCAAATTCAGCATGCTGTTTCAAGAGAGTTCAGTTTACACCCGCGCTTGGATCAGC 2311
2288 TCTCAGGTTCAATGGGAATTCATTTCACTCTGAAGATGAAATTTGAATGAAACCAAA 2347
2312 TCTCAGGTTTCAATGGGAATTCATTTCACTCTGAAGATGAAATTTGAATGAAACCAAA 2371
2348 AGAGCTGCGCAGAGAGAGGAGAGATTTGAACTGCTGAGCTTTGAAAGCTCTCTTT 2407
2372 AAAGGCTG-----GAAGAGAGAGGAGGACTTGAATGTGCTTACATTTGAAGATCTCTTT 2425
2408 GCTTTCGCTACCAAGTGGCCAAAGGATGGAATTCCTGAGTTCAGTCTGCTGTCTCCACA 2467
2426 GCTTTCGATATCAAGTTGCCAAGGAATGGAATTTCTGGAATTTAAGTCTGTTGTTTACA 2485
2468 GAGACTTGGCAGCCAGGAATGTTGGTCAACCCAGGGAGGTTGGTGAAGATCTGTGACT 2527
2486 GAGACTTGGCAGCCAGGAATGTTGGTCAACCCAGGGAGGTTGGTGAAGATCTGTGACT 2545
2528 TTGACTGGCCCGAGACATCTTGAGGACTCCAGCTACGTCAGGGGCAACGACGCGC 2587

2546 TTGGATTGGCTCGAGATATCATGAGTGATTCACATATGTGTGAGGGGCAATCCCGTTC 2605
2588 TGCCGGTGAAGTGGATGGCACCCGAGAGCTTATTTGAAGGATCTACACAAATCAAGAGTG 2647
2606 TGCTCTGAAATGGATGGCCCCCGAAAGCTGTTTGAAGCACTACACCAATTAAGAGTG 2665
2648 ACCTCTGGTCTTACGGCATCTTCTCTGGGAGATATTTTCACTGGGTGTGAACCTTACC 2707
2666 ATGCTGGTCTATGGAATATTAATCTGGGAAATCTTCTCACTGGTGTGAATCTTACC 2725
2708 CTGSCATTCCTGTCGACGCTAACTCTATAACTGATTTCAGAGTGGATTTAAATGAGC 2767
2726 CTGSCATTCCTGTCGACGCTAACTCTTCAAACTGATTTCAAAATGGAATTTAAAATGATC 2785
2768 AGCCATTCCTATGTCACAGAGGATATATCTTGTAAATGCAATCTGCTGGGCTTTTCACT 2827
2786 AGCCATTCCTATGTCACAGAGGATATATCTTGTAAATGCAATCTGCTGGGCTTTTCACT 2845
2828 CAAGGAAGCGGCGCATCTTCCCAACCTGACTTTCAATTTTAGGATGTGAGTGGCAGAGG 2887
2846 CAAGGAAGCGGCGCATCTTCCCTTAAATTTGACTTTCGTTTTTAGGATGTGAGTGGCAGATG 2905
2888 CAGAAGAAG-----CATGTATCAGAAACATCTCATCTCAATCTTACC 2924
2906 CAGAAGAAGCGATGTATCAGAAATGTGGATGGCGGTGTTTCGGAATGTCTCTCACACTTACC 2965
2925 AAAACAGCGCGCGCTCAGCAGAGAGCGGGCTCAGAGC--CCAGTCCGACACGCGCCAGG 2983
2966 AAAACAGCGCGCGCTTTCAGCAGAGAGATGGAATTTGGGGCTACTCTCTCCGAGGCTCAGG 3025
2984 T-GAAGATTCACAGAGAAAGAGTTAGCAGAGAGCGCTTGGACCCCGCC-----ACCCTAG 3038
3026 TCGAAGATTCGTAGAGGAACAAATTTAGTTTAAAGGACTTTCATCCCTCCACTATCCCTAA 3085
3039 CAGGCTGTAGAGCGCAGAGCGCAAGATTAGCTCGCTCT--GAGGAAGCGCCCTACAGCG 3096
3086 CAGGCTGTAGATTACCAAAACCAAGATTAATTTTCATCACTAAAGAAAAATCTATTATCAAC 3145
3097 CGTTGCTTCCGCTGGACTTTTCTCTAGATGCTGTCTGCCATTACTC-----CAAAAGTGA 3149
3146 TCGTCTTCCAGCAGACTTTTCTTAGAGCGGTCTCGGTTTACTCTTTGTTTTCAAAAGGGA 3205
3150 CTTCTATAAAATCAAACTCTCTCTGACAGCGGGAGAGCCAAATAATGAGACTTGTGG 3209
3206 CTTTGTGAAATCAAAATCATCTCTGTCACAAAGCAGCAGGAGCTGATAAGAACTTTATTG 3265
3210 TGAGCCCGCTTACCCTGGGGGCTTTTCCAGAGCTTTGAGGGGAAAGCCATGTATCTCAAA 3269
3266 GAGCATTTGATCTGCAATCCAAAGGCTTCTCAGCGCGCTTGTGAGTGAATTTGTATACCTGAAG 3325
3270 TATAGTATATCTTCTGTAATACGTGAAACAAACCAACCCGTTTTTTTGTAGGGGAAAGC 3329
3326 TACAGTATATCTTCTGTAATACATAAAACAAA-----AGCATTTTGTAGGAGAAAGC 3378
3330 TAAATATGATTTTAAAAATCTATGTTTAAAAATCTATGTAACCTTTTTCATCTATTTAG 3389
3379 TAAATGATTTT--AAGTCTATGTTTAAATATATGTAATTTTTCAGCTATTAG 3436
3390 TGATATATTTATGATGGAATAAACTTTTCTGTAATAAAAAAATAAAAAAATAAAAAA 3449
3437 TGATATATTTATGGTGGGAATAAAATTTCTTACTACAGAAAAAATAAAAAAATAAAAAA 3496
3450 AAAA 3453
3497 AAAA 3500

Search completed: August 27, 2003, 16:01:32
Job time : 1042.85 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 28, 2003, 01:02:19 ; Search time 1007.5 Seconds
(without alignments)
7863.095 Million cell updates/sec

Title: US-09-919-408A-1
Perfect score: 3453
Sequence: 1 GCGGCTGGCTACGGCGC.....AAAAAAAAAAAAAAAAAAAA 3453

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 1533700 seqs, 1147125425 residues

Total number of hits satisfying chosen parameters: 3067400

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:

- 1: /cgn2_6/ptodata/1/pubna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/1/pubna/PCT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/1/pubna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/1/pubna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/1/pubna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/1/pubna/PCTUS_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/1/pubna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/1/pubna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/1/pubna/US09A_PUBCOMB.seq.*
- 10: /cgn2_6/ptodata/1/pubna/US09B_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/1/pubna/US09C_PUBCOMB.seq.*
- 12: /cgn2_6/ptodata/1/pubna/US09_NEW_PUB.seq.*
- 13: /cgn2_6/ptodata/1/pubna/US10A_PUBCOMB.seq.*
- 14: /cgn2_6/ptodata/1/pubna/US10B_PUBCOMB.seq.*
- 15: /cgn2_6/ptodata/1/pubna/US10_NEW_PUB.seq.*
- 16: /cgn2_6/ptodata/1/pubna/US60_NEW_PUB.seq.*
- 17: /cgn2_6/ptodata/1/pubna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	ID	Description
1	3453	100.0	3453	9	US-09-919-408-1	Sequence 1, Appli
2	3453	100.0	3453	10	US-09-872-136-1	Sequence 1, Appli
3	2227.2	64.5	3501	9	US-09-919-408-3	Sequence 3, Appli
4	2227.2	64.5	3501	10	US-09-872-136-3	Sequence 22, Appli
5	2081.8	60.3	3120	10	US-09-982-610-22	Sequence 7538, Ap
6	466.6	13.5	645	10	US-09-796-692-7538	Sequence 7538, Ap
7	466.6	13.5	645	14	US-10-040-862-7538	Sequence 8085, Ap
8	456.4	13.2	610	10	US-09-796-692-8085	Sequence 8085, Ap
9	456.4	13.2	610	14	US-10-040-862-8085	Sequence 8927, Ap
10	454.6	13.2	605	10	US-09-796-692-8927	Sequence 8927, Ap
11	454.6	13.2	605	14	US-10-040-862-8927	Sequence 7976, Ap
12	448.8	13.0	597	10	US-09-796-692-7976	Sequence 7976, Ap
13	448.8	13.0	597	14	US-10-040-862-7976	Sequence 8273, Ap
14	448	13.0	601	10	US-09-796-692-448	Sequence 8273, Ap
15	448	13.0	601	14	US-10-040-862-8273	Sequence 8273, Ap
16	448	13.0	608	10	US-09-796-692-8132	Sequence 8132, Ap

C	17	448	13.0	608	14	US-10-040-862-8132	Sequence 8132, Ap
	18	447.2	13.0	597	10	US-09-796-692-9010	Sequence 9010, Ap
	19	447.2	13.0	597	14	US-10-040-862-9010	Sequence 9010, Ap
	20	440	12.7	605	10	US-09-796-692-8047	Sequence 8047, Ap
	21	440	12.7	605	14	US-10-040-862-8047	Sequence 8047, Ap
	22	439.8	12.7	589	10	US-09-796-692-8630	Sequence 8630, Ap
	23	439.8	12.7	589	14	US-10-040-862-8630	Sequence 8630, Ap
	24	424.4	12.3	594	10	US-09-796-692-8135	Sequence 8135, Ap
	25	424.4	12.3	594	14	US-10-040-862-8135	Sequence 8135, Ap
	26	423.2	12.3	564	10	US-09-796-692-9025	Sequence 9025, Ap
	27	423.2	12.3	564	14	US-10-040-862-9025	Sequence 9025, Ap
	28	416.8	12.1	556	10	US-09-796-692-7883	Sequence 7883, Ap
	29	416.8	12.1	556	14	US-10-040-862-7883	Sequence 7883, Ap
	30	407.2	11.8	550	10	US-09-796-692-8494	Sequence 8494, Ap
	31	407.2	11.8	550	14	US-10-040-862-8494	Sequence 8494, Ap
C	32	399.8	11.6	539	10	US-09-796-692-8255	Sequence 8255, Ap
C	33	399.8	11.6	539	14	US-10-040-862-8255	Sequence 8255, Ap
	34	386.8	11.2	510	10	US-09-796-692-8268	Sequence 8268, Ap
	35	386.8	11.2	510	14	US-10-040-862-8268	Sequence 8268, Ap
	36	360.2	10.4	497	10	US-09-796-692-8408	Sequence 8408, Ap
	37	360.2	10.4	497	14	US-10-040-862-8408	Sequence 8408, Ap
	38	357.4	10.4	3992	10	US-09-944-807-9	Sequence 9, Appli
	39	357.4	10.4	3992	12	US-10-007-326A-292	Sequence 292, App
	40	357.4	10.4	4008	12	US-10-101-510-544	Sequence 544, App
	41	350	10.1	606	10	US-09-796-692-8114	Sequence 8114, Ap
	42	350	10.1	606	14	US-10-040-862-8114	Sequence 8114, Ap
	43	342.2	9.9	508	10	US-09-796-692-8032	Sequence 8032, Ap
	44	342.2	9.9	508	14	US-10-040-862-8032	Sequence 8032, Ap
	45	329	9.5	5084	10	US-09-967-768A-277	Sequence 277, App

ALIGNMENTS

RESULT 1

US-09-919-408-1
; Sequence 1, Application US/09919408
; Patent No. US20020072077A1
; GENERAL INFORMATION:
; APPLICANT: Lemischka, Ihor R.
; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL RECEPTORS AND THEIR LIGANDS
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Imclone Systems Incorporated
; STREET: 180 Varick Street
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10014
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/09/919,408
; FILING DATE: 31-Jul-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/977,451
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US 07/906,397
; FILING DATE: 26-JUN-1992
; APPLICATION NUMBER: US PCT/US92/05401
; FILING DATE: 26-JUN-1992
; APPLICATION NUMBER: TW 81102961
; FILING DATE: 15-APR-1992
; APPLICATION NUMBER: US PCT/US92/02750
; FILING DATE: 02-APR-1992
; APPLICATION NUMBER: US 07/813,593
; FILING DATE: 24-DEC-1991
; APPLICATION NUMBER: US 07/793,065

/ FILING DATE: 15-NOV-1991
/ APPLICATION NUMBER: US 07/728,913
/ FILING DATE: 28-JUN-1991
/ APPLICATION NUMBER: US 07/679,666
/ FILING DATE: 02-APR-1991
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Feit, Irving N.
/ REGISTRATION NUMBER: 28,601
/ REFERENCE/DOCKET NUMBER: LEM-3-7P
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 212-645-1405
/ TELEFAX: 212-645-2054
/ INFORMATION FOR SEQ ID NO: 1:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 3453 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: double
/ TOPOLOGY: linear
/ MOLECULE TYPE: cDNA
/ HYPOTHEetical: NO
/ ANTI-SENSE: NO
/ FRAGMENT TYPE: N-terminal
/ FEATURE:
/ NAME/KEY: mat_peptide
/ LOCATION: 112..3006
/ FEATURE:
/ NAME/KEY: sig_peptide
/ LOCATION: 31..111
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: 31..3009
/ SEQUENCE DESCRIPTION: SEQ ID NO: 1:
/ US-09-919-408-1

Query Match 100.0%; Score 3453; DB 9; Length 3453;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3453; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGGGCTGGCTACCGGCGCTCGGAGGCCATCGGGCGTTGGCGCAGCGACCGG 60
DB 1 GGGGCTGGCTACCGGCGCTCGGAGGCCATCGGGCGTTGGCGCAGCGACCGG 60
QY 61 CGGCTCGTCTGCTGTTGTCAGTAATGATTTGAGACCGTTTACAAACCCAGAC 120
DB 61 CGGCTCGTCTGCTGTTGTCAGTAATGATTTGAGACCGTTTACAAACCCAGAC 120
QY 121 CTGCTCTGATCAAGTGTGTTTAAATCAGTCATGAGAACTTGGCTCATCAGCGGAAAG 180
DB 121 CTGCTCTGATCAAGTGTGTTTAAATCAGTCATGAGAACTTGGCTCATCAGCGGAAAG 180
QY 181 CCATCATCGTACCGAATGGTGCAGGATCCCAAGAGACCTCCAGTGTACCCCGAGCGC 240
DB 181 CCATCATCGTACCGAATGGTGCAGGATCCCAAGAGACCTCCAGTGTACCCCGAGCGC 240
QY 241 CAGAGTGAAGGACGCTATGAAAGCGGCGCCAGCGTGGAGTGGCGGCTCCATC 300
DB 241 CAGAGTGAAGGACGCTATGAAAGCGGCGCCAGCGTGGAGTGGCGGCTCCATC 300
QY 301 ACCCTGCAAGTGCAGTCCGACCCAGGAGACCTTTCTGCTCTGGGTCTTTAAGCAC 360
DB 301 ACCCTGCAAGTGCAGTCCGACCCAGGAGACCTTTCTGCTCTGGGTCTTTAAGCAC 360
QY 361 AGCTCCCTGGGTGCGAGCGCATTTGATTTAAACAAAGAGAAATCGTTTCCATGGCC 420
DB 361 AGCTCCCTGGGTGCGAGCGCATTTGATTTAAACAAAGAGAAATCGTTTCCATGGCC 420
QY 421 ATCTTGAACGTGACAGAGACCCAGGAGGAGATACCTACTCATATTAGCGGAAACGC 480
DB 421 ATCTTGAACGTGACAGAGACCCAGGAGGAGATACCTACTCATATTAGCGGAAACGC 480
QY 481 GCCAACTACACAGTACTGTTTCAAGTGAATGTAAGAGATACACAGCTGTATGCTAAGG 540
DB 481 GCCAACTACACAGTACTGTTTCAAGTGAATGTAAGAGATACACAGCTGTATGCTAAGG 540

QY 541 AGACCTTACTTTAGGAAGATGGAACCAAGATGCATGCTCTGCAATCTCCGAGGCTTT 600
DB 541 AGACCTTACTTTAGGAAGATGGAACCAAGATGCATGCTCTGCAATCTCCGAGGCTTT 600
QY 601 CCGGAGCCCACTGTGGAGTGGTCTCTGAGCTCCCAAGGAAAGCTGTAAAGAGAA 660
DB 601 CCGGAGCCCACTGTGGAGTGGTCTCTGAGCTCCCAAGGAAAGCTGTAAAGAGAA 660
QY 661 GGGCCTGCTGTTGTAGAAAGGAGAAAGGTACTTTCATGAGTTGTTTCGGAACAGATC 720
DB 661 GGGCCTGCTGTTGTAGAAAGGAGAAAGGTACTTTCATGAGTTGTTTCGGAACAGATC 720
QY 721 AGATGCTGCTAGAAATGCACTGGGCGCGAATGCAACCAAGCTGTTCACCATAGATCTA 780
DB 721 AGATGCTGCTAGAAATGCACTGGGCGCGAATGCAACCAAGCTGTTCACCATAGATCTA 780
QY 781 AACGAGCTCTCAGAGCACACTGCCCACTTTCCTGAAAGTGGGGACCCCTTGG 840
DB 781 AACGAGCTCTCAGAGCACACTGCCCACTTTCCTGAAAGTGGGGACCCCTTGG 840
QY 841 ATCAGGTGTAAGGCCATCCATGTGAACCATGGAATTCGGGCTCACCTGGAGCTGAAGAC 900
DB 841 ATCAGGTGTAAGGCCATCCATGTGAACCATGGAATTCGGGCTCACCTGGAGCTGAAGAC 900
QY 901 AAAGCCCTGGAGGAGGCGAGCTACTTTGAGTACCTTCCACAAACAGGACCATG 960
DB 901 AAAGCCCTGGAGGAGGCGAGCTACTTTGAGTACCTTCCACAAACAGGACCATG 960
QY 961 ATTCCGATTTCTTTGGCTTTGTTCTTCCTGGGAGGAAACGACACCGGATATTACCC 1020
DB 961 ATTCCGATTTCTTTGGCTTTGTTCTTCCTGGGAGGAAACGACACCGGATATTACCC 1020
QY 1021 TGCTCTTCTCAAGCACCCAGCCAGTACAGCTGTGTGACCATCTCTAGAAAAGGTTT 1080
DB 1021 TGCTCTTCTCAAGCACCCAGCCAGTACAGCTGTGTGACCATCTCTAGAAAAGGTTT 1080
QY 1081 ATAAACGCTACCAAGCTCGCAAGAGAGTATGAAATTTGACCCGTCGAAAAGTTCTGCTTC 1140
DB 1081 ATAAACGCTACCAAGCTCGCAAGAGAGTATGAAATTTGACCCGTCGAAAAGTTCTGCTTC 1140
QY 1141 TCAGTCAGGTTTAAAGCTACCCAGTAACCAAGTACCGATCCGATGCGGATCTTCTCAAGCTCA 1200
DB 1141 TCAGTCAGGTTTAAAGCTACCCAGTAACCAAGTACCGATCCGATGCGGATCTTCTCAAGCTCA 1200
QY 1201 TTTCTCTGTGAACAGAGAGGCTTGGAGGATGGGTACAGCATATCTAAATTTTGGATCAT 1260
DB 1201 TTTCTCTGTGAACAGAGAGGCTTGGAGGATGGGTACAGCATATCTAAATTTTGGATCAT 1260
QY 1261 AAGAACAGCCAGGAGTACATATTTCTATGCAAGAAATGATGACGCCCAAGTTTCAACAAA 1320
DB 1261 AAGAACAGCCAGGAGTACATATTTCTATGCAAGAAATGATGACGCCCAAGTTTCAACAAA 1320
QY 1321 ATGTTTCAGCTGAATATAGAAAGAAACCTCAAGTCTAGCAAAATGCTCAGCCAGCCAG 1380
DB 1321 ATGTTTCAGCTGAATATAGAAAGAAACCTCAAGTCTAGCAAAATGCTCAGCCAGCCAG 1380
QY 1381 GGGTCTGTTCTCTGATGGCTACCGCTACCTCTTGGACCTTGGAGAGTGTTCGGAC 1440
DB 1381 GGGTCTGTTCTCTGATGGCTACCGCTACCTCTTGGACCTTGGAGAGTGTTCGGAC 1440
QY 1441 AAATCTCCAAATGCAAGGAGAAATCCCAAGAGAGTTTGGAAATAAAAGGCTAAACAGA 1500
DB 1441 AAATCTCCAAATGCAAGGAGAAATCCCAAGAGAGTTTGGAAATAAAAGGCTAAACAGA 1500
QY 1501 AAAGTGTGTTGGCAGTGGGTGTCGAGCAGTACTCTAAATATGAGTGAAGCGGAAAGG 1560
DB 1501 AAAGTGTGTTGGCAGTGGGTGTCGAGCAGTACTCTAAATATGAGTGAAGCGGAAAGG 1560
QY 1561 CTTCCTGGTCAAAATGCTGTCGCTACAAATTTATGGCACGCTTTCGAAACCATCTTTTAA 1620
DB 1561 CTTCCTGGTCAAAATGCTGTCGCTACAAATTTATGGCACGCTTTCGAAACCATCTTTTAA 1620

1621 AACTCAGGCGCCCTTCCCTTTCATCCAGACAACTCTCTTCTATGCGACCATTTGG 1680
1621 AACTCAGGCGCCCTTCCCTTTCATCCAGACAACTCTCTTCTATGCGACCATTTGG 1680
1681 CTCGTCTCCCTTCAATGTTGTTCTCATTTGTTGATCTGCCACAAATACAAAAGCAA 1740
1681 CTCGTCTCCCTTCAATGTTGTTCTCATTTGTTGATCTGCCACAAATACAAAAGCAA 1740
1741 TTTAGGTACGAGAGTCACTGATCCAGGTGATCCAGGTGATCCAGGTGATCCAGGTGAT 1800
1741 TTTAGGTACGAGAGTCACTGATCCAGGTGATCCAGGTGATCCAGGTGATCCAGGTGAT 1800
1801 TTTAGGTACGAGAGTCACTGATCCAGGTGATCCAGGTGATCCAGGTGATCCAGGTGAT 1860
1801 TTTAGGTACGAGAGTCACTGATCCAGGTGATCCAGGTGATCCAGGTGATCCAGGTGAT 1860
1861 TTTAGGTACGAGAGTCACTGATCCAGGTGATCCAGGTGATCCAGGTGATCCAGGTGAT 1920
1861 TTTAGGTACGAGAGTCACTGATCCAGGTGATCCAGGTGATCCAGGTGATCCAGGTGAT 1920
1921 TTTAGGTACGAGAGTCACTGATCCAGGTGATCCAGGTGATCCAGGTGATCCAGGTGAT 1980
1921 TTTAGGTACGAGAGTCACTGATCCAGGTGATCCAGGTGATCCAGGTGATCCAGGTGAT 1980
1981 TTTAGGTACGAGAGTCACTGATCCAGGTGATCCAGGTGATCCAGGTGATCCAGGTGAT 2040
1981 TTTAGGTACGAGAGTCACTGATCCAGGTGATCCAGGTGATCCAGGTGATCCAGGTGAT 2040
2041 TTTAGGTACGAGAGTCACTGATCCAGGTGATCCAGGTGATCCAGGTGATCCAGGTGAT 2100
2041 TTTAGGTACGAGAGTCACTGATCCAGGTGATCCAGGTGATCCAGGTGATCCAGGTGAT 2100
2101 TTTAGGTACGAGAGTCACTGATCCAGGTGATCCAGGTGATCCAGGTGATCCAGGTGAT 2160
2101 TTTAGGTACGAGAGTCACTGATCCAGGTGATCCAGGTGATCCAGGTGATCCAGGTGAT 2160
2161 TTTAGGTACGAGAGTCACTGATCCAGGTGATCCAGGTGATCCAGGTGATCCAGGTGAT 2220
2161 TTTAGGTACGAGAGTCACTGATCCAGGTGATCCAGGTGATCCAGGTGATCCAGGTGAT 2220
2221 TTTAGGTACGAGAGTCACTGATCCAGGTGATCCAGGTGATCCAGGTGATCCAGGTGAT 2280
2221 TTTAGGTACGAGAGTCACTGATCCAGGTGATCCAGGTGATCCAGGTGATCCAGGTGAT 2280
2281 TTTAGGTACGAGAGTCACTGATCCAGGTGATCCAGGTGATCCAGGTGATCCAGGTGAT 2340
2281 TTTAGGTACGAGAGTCACTGATCCAGGTGATCCAGGTGATCCAGGTGATCCAGGTGAT 2340
2341 TTTAGGTACGAGAGTCACTGATCCAGGTGATCCAGGTGATCCAGGTGATCCAGGTGAT 2400
2341 TTTAGGTACGAGAGTCACTGATCCAGGTGATCCAGGTGATCCAGGTGATCCAGGTGAT 2400
2401 TTTAGGTACGAGAGTCACTGATCCAGGTGATCCAGGTGATCCAGGTGATCCAGGTGAT 2460
2401 TTTAGGTACGAGAGTCACTGATCCAGGTGATCCAGGTGATCCAGGTGATCCAGGTGAT 2460
2461 TTTAGGTACGAGAGTCACTGATCCAGGTGATCCAGGTGATCCAGGTGATCCAGGTGAT 2520
2461 TTTAGGTACGAGAGTCACTGATCCAGGTGATCCAGGTGATCCAGGTGATCCAGGTGAT 2520
2521 TTTAGGTACGAGAGTCACTGATCCAGGTGATCCAGGTGATCCAGGTGATCCAGGTGAT 2580
2521 TTTAGGTACGAGAGTCACTGATCCAGGTGATCCAGGTGATCCAGGTGATCCAGGTGAT 2580
2581 TTTAGGTACGAGAGTCACTGATCCAGGTGATCCAGGTGATCCAGGTGATCCAGGTGAT 2640
2581 TTTAGGTACGAGAGTCACTGATCCAGGTGATCCAGGTGATCCAGGTGATCCAGGTGAT 2640
2641 TTTAGGTACGAGAGTCACTGATCCAGGTGATCCAGGTGATCCAGGTGATCCAGGTGAT 2700
2641 TTTAGGTACGAGAGTCACTGATCCAGGTGATCCAGGTGATCCAGGTGATCCAGGTGAT 2700
2701 TTTAGGTACGAGAGTCACTGATCCAGGTGATCCAGGTGATCCAGGTGATCCAGGTGAT 2760

2701 TTTAGGTACGAGAGTCACTGATCCAGGTGATCCAGGTGATCCAGGTGATCCAGGTGAT 2760
2761 TTTAGGTACGAGAGTCACTGATCCAGGTGATCCAGGTGATCCAGGTGATCCAGGTGAT 2820
2761 TTTAGGTACGAGAGTCACTGATCCAGGTGATCCAGGTGATCCAGGTGATCCAGGTGAT 2820
2821 TTTAGGTACGAGAGTCACTGATCCAGGTGATCCAGGTGATCCAGGTGATCCAGGTGAT 2880
2821 TTTAGGTACGAGAGTCACTGATCCAGGTGATCCAGGTGATCCAGGTGATCCAGGTGAT 2880
2881 TTTAGGTACGAGAGTCACTGATCCAGGTGATCCAGGTGATCCAGGTGATCCAGGTGAT 2940
2881 TTTAGGTACGAGAGTCACTGATCCAGGTGATCCAGGTGATCCAGGTGATCCAGGTGAT 2940
2941 TTTAGGTACGAGAGTCACTGATCCAGGTGATCCAGGTGATCCAGGTGATCCAGGTGAT 3000
2941 TTTAGGTACGAGAGTCACTGATCCAGGTGATCCAGGTGATCCAGGTGATCCAGGTGAT 3000
3001 TTTAGGTACGAGAGTCACTGATCCAGGTGATCCAGGTGATCCAGGTGATCCAGGTGAT 3060
3001 TTTAGGTACGAGAGTCACTGATCCAGGTGATCCAGGTGATCCAGGTGATCCAGGTGAT 3060
3061 TTTAGGTACGAGAGTCACTGATCCAGGTGATCCAGGTGATCCAGGTGATCCAGGTGAT 3120
3061 TTTAGGTACGAGAGTCACTGATCCAGGTGATCCAGGTGATCCAGGTGATCCAGGTGAT 3120
3121 TTTAGGTACGAGAGTCACTGATCCAGGTGATCCAGGTGATCCAGGTGATCCAGGTGAT 3180
3121 TTTAGGTACGAGAGTCACTGATCCAGGTGATCCAGGTGATCCAGGTGATCCAGGTGAT 3180
3181 TTTAGGTACGAGAGTCACTGATCCAGGTGATCCAGGTGATCCAGGTGATCCAGGTGAT 3240
3181 TTTAGGTACGAGAGTCACTGATCCAGGTGATCCAGGTGATCCAGGTGATCCAGGTGAT 3240
3241 TTTAGGTACGAGAGTCACTGATCCAGGTGATCCAGGTGATCCAGGTGATCCAGGTGAT 3300
3241 TTTAGGTACGAGAGTCACTGATCCAGGTGATCCAGGTGATCCAGGTGATCCAGGTGAT 3300
3301 TTTAGGTACGAGAGTCACTGATCCAGGTGATCCAGGTGATCCAGGTGATCCAGGTGAT 3360
3301 TTTAGGTACGAGAGTCACTGATCCAGGTGATCCAGGTGATCCAGGTGATCCAGGTGAT 3360
3361 TTTAGGTACGAGAGTCACTGATCCAGGTGATCCAGGTGATCCAGGTGATCCAGGTGAT 3420
3361 TTTAGGTACGAGAGTCACTGATCCAGGTGATCCAGGTGATCCAGGTGATCCAGGTGAT 3420
3421 TTTAGGTACGAGAGTCACTGATCCAGGTGATCCAGGTGATCCAGGTGATCCAGGTGAT 3453
3421 TTTAGGTACGAGAGTCACTGATCCAGGTGATCCAGGTGATCCAGGTGATCCAGGTGAT 3453

RESULT 2

US-09-872-136-1
; Sequence 1, Application US/09872136
; Patent No. US20020119545A1
; GENERAL INFORMATION:
; APPLICANT: Lemischka, Ihor R.
; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL
; RECEPTORS AND THEIR LIGANDS
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Imclone Systems Incorporated
; STREET: 180 Varick Street
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10014
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25

Db 1261 AAGAAACAGCCAGGAGTACATATTCTATGCAGAAAATGATGACGCCAGTTTCAACAAA 1320
 Qy 1321 ATGTTACGCTGAATATAAGAAAGAAACCTCAAGTGTAGCAAAATGCTCAGCCAGCCAG 1380
 Db 1321 ATGTTACGCTGAATATAAGAAAGAAACCTCAAGTGTAGCAAAATGCTCAGCCAGCCAG 1380
 Qy 1381 GCGTCTGTTCTCTGATGCTACCCGCTACCCCTCTTGGACCTTGGAAAGAGTGTTCGGAC 1440
 Db 1381 GCGTCTGTTCTCTGATGCTACCCGCTACCCCTCTTGGACCTTGGAAAGAGTGTTCGGAC 1440
 Qy 1441 AAATCTCCCAATTCACGAGGAAATCCAGAGGAGTTTGGAAATAAAAAAGCTTAACAGA 1500
 Db 1441 AAATCTCCCAATTCACGAGGAAATCCAGAGGAGTTTGGAAATAAAAAAGCTTAACAGA 1500
 Qy 1501 AAAGTGTTCGCGAGTGGGTGTGAGAGATCTCTAAATATGAGTGAGGCCGCGGAAGGG 1560
 Db 1501 AAAGTGTTCGCGAGTGGGTGTGAGAGATCTCTAAATATGAGTGAGGCCGCGGAAGGG 1560
 Qy 1561 CTTCTGCTCAATGCTGTGCTCAATCTATGCGCAGCTTTCGGAACCATCTTTTAA 1620
 Db 1561 CTTCTGCTCAATGCTGTGCTCAATCTATGCGCAGCTTTCGGAACCATCTTTTAA 1620
 Qy 1621 AACTCACAGCCGCTTCCCTTTTCATCAAGACAACTCTCTTCTATGCGACCATTTGGG 1680
 Db 1621 AACTCACAGCCGCTTCCCTTTTCATCAAGACAACTCTCTTCTATGCGACCATTTGGG 1680
 Qy 1681 CTCTGCTCTCCCTTCATTTGTTGTTCTCATTTGTTGATCTGCCAATAACAAAAGCAA 1740
 Db 1681 CTCTGCTCTCCCTTCATTTGTTGTTCTCATTTGTTGATCTGCCAATAACAAAAGCAA 1740
 Qy 1741 TTTAGGTACGAGAGTCACTGATGATCCAGGTGACTGCGCCCTTGGATAACGAGTAC 1800
 Db 1741 TTTAGGTACGAGAGTCACTGATGATCCAGGTGACTGCGCCCTTGGATAACGAGTAC 1800
 Qy 1801 TTCTACGTTGACTTCAGGAGCTATGAATATGACCTTAAGTGGAGTTCCTCCAGAGAGAAC 1860
 Db 1801 TTCTACGTTGACTTCAGGAGCTATGAATATGACCTTAAGTGGAGTTCCTCCAGAGAGAAC 1860
 Qy 1861 TTAGAGTTTGGGAAGTCTCTGGGTCTGGCGCTTTCGGGAGGTGATGAACGCCACGCC 1920
 Db 1861 TTAGAGTTTGGGAAGTCTCTGGGTCTGGCGCTTTCGGGAGGTGATGAACGCCACGCC 1920
 Qy 1921 TATGGCAATAGTAAACCGGAGTCTCAATTCAGTGGCGGTGAAGATGCTAAAGAGAAA 1980
 Db 1921 TATGGCAATAGTAAACCGGAGTCTCAATTCAGTGGCGGTGAAGATGCTAAAGAGAAA 1980
 Qy 1981 GCTGACAGCTGTGAAAAGAGCTCTCATGTCGGAGCTCAAAATGATGACCCACCTGGGA 2040
 Db 1981 GCTGACAGCTGTGAAAAGAGCTCTCATGTCGGAGCTCAAAATGATGACCCACCTGGGA 2040
 Qy 2041 CACCATGACAACTGCTGAATCTGCTGGGGGATGACACATGTCAGGGCCAGTGTACTTG 2100
 Db 2041 CACCATGACAACTGCTGAATCTGCTGGGGGATGACACATGTCAGGGCCAGTGTACTTG 2100
 Qy 2101 ATTTTGAATATGTTGCTATGTTGCTCTCAACTCTCAAGTGAAGTAAAGAGAGAG 2160
 Db 2101 ATTTTGAATATGTTGCTATGTTGCTCTCAACTCTCAAGTGAAGTAAAGAGAGAG 2160
 Qy 2161 TTTTACAGGACATGACAGAGATTTTAAAGNACATTAATTCAGTCTTACCCCTACTTTC 2220
 Db 2161 TTTTACAGGACATGACAGAGATTTTAAAGNACATTAATTCAGTCTTACCCCTACTTTC 2220
 Qy 2221 CAGGCACATTCAAATTCAGAGATGCTGTTTACGAGAGATTCAGTTACACCCGCCCTTG 2280
 Db 2221 CAGGCACATTCAAATTCAGAGATGCTGTTTACGAGAGATTCAGTTACACCCGCCCTTG 2280
 Qy 2281 GATCAGCTCTCAGGGTTCATTTGGGAATTCATTCATTCGAGAGATGAGATTGAATGAA 2340
 Db 2281 GATCAGCTCTCAGGGTTCATTTGGGAATTCATTCATTCGAGAGATGAGATTGAATGAA 2340
 Qy 2341 AACGGAAGAGCTGGCAGAGAGAGAGAGAGATTTGAACTGCTGACGTTTGAAGAC 2400
 Db 2341 AACGGAAGAGCTGGCAGAGAGAGAGAGAGATTTGAACTGCTGACGTTTGAAGAC 2400

Qy 2401 CTCCTTTGCTTTGGTACCAGTGGCCAAAGGATGGAATTCCTGGAGTTCAGTCTGTGT 2460
 Db 2401 CTCCTTTGCTTTGGTACCAGTGGCCAAAGGATGGAATTCCTGGAGTTCAGTCTGTGT 2460
 Qy 2461 GTCCACAGAGACTCTGGCAGCAGGAATGTGTGTGTACCCACCGGAAGGTGTGAAGATC 2520
 Db 2461 GTCCACAGAGACTCTGGCAGCAGGAATGTGTGTGTACCCACCGGAAGGTGTGAAGATC 2520
 Qy 2521 TGTGACTTTGGACTGGCCCGGAGACATCTGTAGCGACTCCAGTACGTCTGTCAGGGGCAAC 2580
 Db 2521 TGTGACTTTGGACTGGCCCGGAGACATCTGTAGCGACTCCAGTACGTCTGTCAGGGGCAAC 2580
 Qy 2581 GCACGGCTGCGCGTGAAGTGGATGGCACCAGAGACTTATTTGAAGGGATCTACACAATC 2640
 Db 2581 GCACGGCTGCGCGTGAAGTGGATGGCACCAGAGACTTATTTGAAGGGATCTACACAATC 2640
 Qy 2641 AAGAGTGAAGTCTGCTCTACCGCATCTCTCTGGGAGATATTTTCACTGGGTGTGAAC 2700
 Db 2641 AAGAGTGAAGTCTGCTCTACCGCATCTCTCTGGGAGATATTTTCACTGGGTGTGAAC 2700
 Qy 2701 CTTTACCTGGCATCTCTGTGCGCTAACTTCTTATAAATGATTCAGAGTGGATTTAAA 2760
 Db 2701 CTTTACCTGGCATCTCTGTGCGCTAACTTCTTATAAATGATTCAGAGTGGATTTAAA 2760
 Qy 2761 ATGAGCAGCCATCTATGCGCAGAGAGGATATACCTTTGTAATGCAATCTCTGCTGGGCT 2820
 Db 2761 ATGAGCAGCCATCTATGCGCAGAGAGGATATACCTTTGTAATGCAATCTCTGCTGGGCT 2820
 Qy 2821 TTTGACTCAAGGAAGCGGCCATCTCTCCCAACCTGACTTCTATTTTATAGGATGTGAGTGT 2880
 Db 2821 TTTGACTCAAGGAAGCGGCCATCTCTCCCAACCTGACTTCTATTTTATAGGATGTGAGTGT 2880
 Qy 2881 GCAGAGCAGAGAGAGAGATGATCAGAAATCATCATCTTACCAAAACAGCGGCCCT 2940
 Db 2881 GCAGAGCAGAGAGAGAGATGATCAGAAATCATCATCTTACCAAAACAGCGGCCCT 2940
 Qy 2941 CAGCAG 3000
 Db 2941 CAGCAG 3000
 Qy 3001 AGAAGTTAGCAG 3060
 Db 3001 AGAAGTTAGCAG 3060
 Qy 3061 AGATTAGCCTCGCTCTGAGAAAGCGCCCTACAGCGCGTGTCTCGTGGAGCTTTTCTCT 3120
 Db 3061 AGATTAGCCTCGCTCTGAGAAAGCGCCCTACAGCGCGTGTCTCGTGGAGCTTTTCTCT 3120
 Qy 3121 AGATTGCTGTGCTGCTTACTCCAAAGTACTTCTATAAATCAAAACCTCTCTCGCACAG 3180
 Db 3121 AGATTGCTGTGCTGCTTACTCCAAAGTACTTCTATAAATCAAAACCTCTCTCGCACAG 3180
 Qy 3181 GCGGAG 3240
 Db 3181 GCGGAG 3240
 Qy 3241 AGCTTTGAGGGGAAAGCCATGTATCTGAAATATAGTATATTTTAAATCTATGTTTAA 3300
 Db 3241 AGCTTTGAGGGGAAAGCCATGTATCTGAAATATAGTATATTTTAAATCTATGTTTAA 3300
 Qy 3301 ACCAAACCCGTTTTTGTCTAAGGAAAGCTAAATATGATTTTAAATCTATGTTTAA 3360
 Db 3301 ACCAAACCCGTTTTTGTCTAAGGAAAGCTAAATATGATTTTAAATCTATGTTTAA 3360
 Qy 3361 AATCTATGTAATCTTCTATCTATTTAGTATATATTTTATGAGTGGAAAAATAACTTTTC 3420
 Db 3361 AATCTATGTAATCTTCTATCTATTTAGTATATATTTTATGAGTGGAAAAATAACTTTTC 3420
 Qy 3421 TACTCTAAAAAAG 3453
 Db 3421 TACTCTAAAAAAG 3453

RESULT 3
US-09-919-408-3
Sequence 3, Application US/09919408
Patent No. US20020072077A1
GENERAL INFORMATION:
APPLICANT: Lemischka, Ihor R.
TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL
RECEPTORS AND THEIR LIGANDS
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Imclone Systems Incorporated
STREET: 180 Varick Street
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10014
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/919,408
FILING DATE: 31-Jul-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/977,451
FILING DATE: <Unknown>
APPLICATION NUMBER: US 07/906,397
FILING DATE: 26-JUN-1992
APPLICATION NUMBER: US PCT/US92/05401
FILING DATE: 26-JUN-1992
APPLICATION NUMBER: TW 81102961
FILING DATE: 15-APR-1992
APPLICATION NUMBER: US PCT/US92/02750
FILING DATE: 02-APR-1992
APPLICATION NUMBER: US 07/813,593
FILING DATE: 24-DEC-1991
APPLICATION NUMBER: US 07/793,065
FILING DATE: 15-NOV-1991
APPLICATION NUMBER: US 07/728,913
FILING DATE: 28-JUN-1991
APPLICATION NUMBER: US 07/679,666
FILING DATE: 02-APR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Feit, Irving N.
REGISTRATION NUMBER: 28,601
REFERENCE/DOCKET NUMBER: LEM-3-7P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-645-1405
TELEFAX: 212-645-2054
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 3501 base pairs
TYPE: nucleic acid
STRAINEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
FEATURE:
NAME/KEY: CDS
LOCATION: 58..3039
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 139..3036
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 58..138
SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-919-408-3

Query Match 64.5%; Score 2227.2; DB 9; Length 3501;
Best Local Similarity 79.6%; Pred. No. 0;
Matches 2775; Conservative 0; Mismatches 653; Indels 56; Gaps 10;
QY 8 GGCTACCGCGCTCCGGAGGCCATCGGGCGTGGGGAGCCAGCGAGCGCGGGGTGC 67
DB 35 GGGGACCCCGGCTCCGGAGGCCATCGGGCGTGGGGAGCCAGCGAGCGCGGGGTGC 91
QY 68 TGCTGCTTGTGTTTGTGTCAGTAATGATTTCTTGAGACCGTTACAAACCAAGACTCGCTG 127
DB 92 CGCTGCTGTTGTTTCTGCAATGATATTGGGACTATTACAAATCAAGATCTGCTG 151
QY 128 TGATCAAGTGTGTTTAAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 187
DB 152 TGATCAAGTGTGTTTAAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 211
QY 188 CGTACCGATGTCGCGAGGATCCCAAGAGACTCCAGTGATCCCCGAGGCGCAGAGTG 247
DB 212 CATATCCCATGGTATCAGAAATCCCGGAAGACTCGGGTGTGGTTCAGACCCAGAGCT 271
QY 248 AAGGACCGGTATATGAAGCGGCCACCGTGGAGTGGCCGAGTCTGGTCCATCACCTGC 307
DB 272 CAGGACAGTGTACGAGCTGCCGTGTGGAAGTGGATGATCTCTCTCCATCACACTGC 331
QY 308 AAGTGCAGCTCGCCACCCAGGGGACCTTTCCTGCTCTGGGTCTTTAAGCACAGCTCCC 367
DB 332 AAGTGCAGTGTGATGCTCCCGAGGAACTTTCTGCTCTGGGTCTTTAAGCACAGCTCCC 391
QY 368 TGGGCTGCGCAGCGCACTTTGATTTACAAAACAGAGGAATCGTTTCCATGGCCATCTGA 427
DB 392 TGAATTGGCCAGCCACATTTGATTTTACAAAACAGAGGAGTTGTTTCCATGGTCAATTTGA 451
QY 428 ACCTGACAGAGACCCAGGAGGAAATACCTACTCCATATTCAGAGCGAACCGCCAACT 487
DB 452 AAATGACAGAAACCCAGCTGGAGAAATACCTACTTTTATTCAGAGTGAAGTCACTCAAT 511
QY 488 ACACAGTACTGTTTCAAGTGAATGTAAGAGATACACAGCTGTATGTGTAAAGAGACCTT 547
DB 512 ACACAAATATTGTTTACAGTGAATGTAAGAAATACCTGCTTTACACATTAAGAGACCTT 571
QY 548 ACTTTAGGAAGATGGAACCCAGGATGCACTGCTCTGCAATCTCCGAGGGTGTCCGAGC 607
DB 572 ACTTTAGAAATGGAACCCAGGAGCGCCCTGCTGCAATCTCGAGAGCGTTCCAGAGC 631
QY 608 CCACTGTGGGTGCTCTGCACTCCCAAGGAAAGCTGTAAGAGAGGAGGCGCCCTG 667
DB 632 CGATCGGAATGGTGTCTTTGCGATTACAGGGGGAAGCTGTAAAGAGAAAGTCCAG 691
QY 668 CTGTTGTCAGAAAGGAGGAAAGGTACTTTTCATGAGTTGTTTCGGAACAGACATCAGATGCT 727
DB 692 CTGTTGTTAAAGAGGAGGAAAGGTCTTCATGAATTAATTTGGGACGACATAAGGTGCT 751
QY 728 GTGCTAGAAATGCACTGGCCGCGGAATGCAACAAGCTGTTCCACATAGATCTAAACAGG 787
DB 752 GTGCCAGAAATGAATGGGAGGGAATGCACAGGCTGTTTCAATAGATCTTAAATCAA 811
QY 788 CTCCTCAGAGCACACTGCCCGAGTTATTCCTGAAAGTGGGGAACCTTTGCGATCAGT 847
DB 812 CTCCTCAGACCAATGCCCAATTAATTTCTTAAAGTAGGGGAACCTTTATGGATAAGGT 871
QY 848 GTAAGGCCATCCATGTGAACCATGGATTTCGGGCTCACCTGGGAGCTGGAAGCAAGGCC 907
DB 872 GCAAGCTGTTCAATGTGAACCATGGATTTCGGGCTCACCTGGGAATTAAGAAACAAAGAC 931
QY 908 TGGAGGAGGAGCAGTACTTTGAGATGATGATCTACTCCACAAACAGAGGACCATGATTCGA 967
DB 932 TCGAGGAGGAGCACTACTTTGAGATGATGATCTACTTCAACAAACAGAGACTATGATACGA 991
QY 968 TTCTCTTGGCTTTTGTGCTCTTCCGTGGAGGAGGAGCAACCGGATATTTACACCTGCTCT 1027
DB 992 TTCTGTTTGTCTTTGTATCATCATGTCAGTGAAGAAACGACACCGGATACTACTACCTTGTCT 1051

QY	1028	CCTCAAAGCACCCAGCAGTCTGGTGGTACCATCTAGAAAAGGGTTTATAAAG	1087	Db	2132	AATACCTGTTGCTATGGTGATCTTCTCAACTATCTAAGAAGTAAAGAGAAAATTTTCA	2191
Db	1052	CTTCAAAGCATCCAGTCAATCAGCTTTGGTTACCATCGTAGAAAAGGGATTTATAAATG	1111	QY	2168	GGACATGGACAGAGATTTTAAAGGAACATAATTTCACTTCTTACCTACTTTTCCAGGCAC	2227
QY	1088	CTACCAAGCTCGCAAGAGATGAAATTCACCGGTACGAAAGTTCTGCTTCTCAGTCA	1147	Db	2192	GGACTTGGACAGAGATTTTCAAGGAACACAATTTCACTTCTTACCCCACTTTTCCAATCAC	2251
Db	1112	CTACCAATTTCAAGTGAAGATTTGAATTTGACCAATATGAAGATTTTGTCTTCTGTC	1171	QY	2228	ATTCAAATTTCCAGCATGCTGCTGTTTCAAGGAAGTTTCACTTACCCGCTTGGATCAGC	2287
QY	1148	GGTTTAAAGCGTACCAACGAATCCGATGCAAGTGGATCTTCTCTCAAGCCTCATTTTCCCT	1207	Db	2252	ATCCAAATTTCCAGCATGCTGCTGTTTCAAGGAAGTTTCAAGATACCCCGGACTCGGATCA	2311
Db	1172	GGTTTAAAGCGTACCAACGAATCCGATGCAAGTGGATCTTCTCTCAAGCCTCATTTTCCCT	1231	QY	2288	TCTCAGGTTTCAATGGGAATTCATTTCAATTTCTGAAGATGAGATTTGAATGAACACCA	2347
QY	1208	GTGAAACAGAGAGCCTGGAGGATGGGTACAGCATATCTAAATTTTGGATCATGAAGACA	1267	Db	2312	TCTCAGGCTTCTATGGGAATTCATTTCACTCTGAAGATGAAATTTGAATGAACACCA	2371
Db	1232	GTGAGCAAAAGGGTCTTGATAACGGATACAGCATATCCAAAGTTTTCGAATTCATAAGCAC	1291	QY	2348	AGAGGCTGGCAGAAAGAGAGAGAGATTTGAACGCTGCTGACGTTTGAAGACCTCTCTTT	2407
QY	1268	AGCCAGGAGTACATATTTCTATGCAAAATGATGACGCCAGTTCACCAAAATGTTCA	1327	Db	2372	AAAGGCTG-----GAAGAGAGAGGAGACTTGAATGTGCTTACATTTGAAGATCTTCTTT	2425
Db	1292	AGCCAGGAGATATATTTTCCATGCGAAGAAATGATGATGCCCAATTTTACCAAAATGTTCA	1351	QY	2408	GCTTTGCTTACCAAGTGGCCAAAGGCATGGAATTTCTGGAGTTTCAAGTCTGTGTCCACA	2467
QY	1328	CGCTGAATATAAGAAAGAAACCTCAAGTGTAGCAAAATGCTCAGCCAGCCAGGCGTCTCT	1387	Db	2426	GCTTTGCTATATCAAGTGGCCAAAGGAATGGAATTTCTGGAATTTTAAAGTCTGTGTTCACA	2485
Db	1352	CGCTGAATATAAGAAAGAAACCTCAAGTGTCTCGAGAGCATCGCAAGTTCAGGCGTCTCT	1411	QY	2468	GAGACCTGGCAGCAGGAATGTGTTGGTCAACCACCGGAAGGTGTGGAAGATCTGTGACT	2527
QY	1388	GTTCCTCTGATGCTACCCGCTACCTCTTGGACCTGGAAGAGTGTTCGGAACAATCTC	1447	Db	2486	GAGACCTGGCCGCGCAGGAACGTGCTTGTCTACCCCACGGGAAGTGTGTGAAGATATGTGACT	2545
Db	1412	GTTCCTCGGATGATACCCATTAACATCTTGGACCTGGAAAGTGTTCAGACAAGTCTC	1471	QY	2528	TTGACCTGGCCCGAGACATCTCTGAGGACTCCAGCTACGCTGTGAGGGGCAACGCAAGC	2587
QY	1448	CCAAATTCAGCGAGGAATCCAGAGAGTGTGGAATTAAGAGCTTAAACAGAAAGTGT	1507	Db	2546	TTGACCTGGCTCGAGATATCATGATGATTTCCAATCTATGTTGTGAGGGGCAATGCCGCTC	2605
Db	1472	CCAACTGCACAGAGAGATCACAGAGGAGTCTGGAATGAAGAGGCTTAAACAGAAAGTGT	1531	QY	2588	TGCGGTGAAGTGGATGGCACCCGAGAGCTTATTTGAAGGATCTTACACAACTCAAGAGTG	2647
QY	1508	TTGGCAGTGGGTCTCGAGCAGTACTCTAAATATGAGTGGCGCGGAAAGGGCTTCTGG	1567	Db	2606	TGCGGTGAAGTGGATGGCCCGGAGAGGCTGTTTGAAGGCACTTACACAACTTAAAGAGTG	2665
Db	1532	TTGACAGTGGGTCTCGAGCAGTACTCTAAATATGAGTGGCGCGGAAAGGGCTTCTGG	1591	QY	2648	ACGTCTGGTCTTACGGCATCTCTCTGGGAGATATTTTCACTGGGTGTGAACCTTTACC	2707
QY	1568	TCAAATCTGTGCGTACAAATCTTATGGGACGCTCTTGGAAACCAATCTTTTAACTCAC	1627	Db	2666	ATGCTGGTCTATGGAATATTTACTGTGGGAATCTTCTCACTTGTGTGAATCTTACC	2725
Db	1592	TCAAGTCTGTGCATACAAATCTTCCCTTGGCAGTCTTGTGAGAGCATCTTTTAACTCTC	1651	QY	2708	CTGGCAATCTCGTTCGACGCTTCTTATATACTGATTTAGATGAGTGTGATTTAAATGGAGC	2767
QY	1628	CAGGCCCTTCCCTTCTATCCAGACACATCTCTCTCATGCGACCATTTGGGCTCTGTC	1687	Db	2726	CTGGCAATCTCGGTTGTGATCTTAACTTCTACAACTGATTTCAAAATGGAATTTAAATGGATC	2785
Db	1652	CAGGCCCTTCCCTTCTATCCAGACACATCTCTCTCATGCAACAAATTTGGTGTGTC	1711	QY	2768	AGCCATTTCTATGCCACAGAGGATATCTTTGTATGCAATCTCTGCTGGGCTTTTGGACT	2827
QY	1688	TCCCTTCTATGTTGTTCTCATTTGTTGATCTGCCCAATAACAAAGCAATTTAGGT	1747	Db	2786	AGCCATTTTATGCTACAGAGAAATATATCATTAATGCAATCTCTGCTGGGCTTTTGGACT	2845
Db	1712	TCTCTTCTATGTTGTTTAACTTCTATTTGTCTCAAGTACAAAGCAATTTAGGT	1771	QY	2828	CAAGAGCGGCCATCTCTTCCCAACCTGACTTCAATTTTATAGGATGTGAGTGTGAGGAGG	2887
QY	1748	ACGAGTCTAGCTCAGATGATCAGGTGATGTCGCCATCTGTCGATGATGATGATGATGATG	1807	Db	2846	CAAGGAAAAGGCCATCTCTTCCCTTAATTTGACTTCTGTTTATAGGATGTGAGTGTGAGGAGG	2905
Db	1772	ATGAAGCCAGCTACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	1831	QY	2888	CAGAAGAG-----CATGTATCAGAAATCCATCCATCTACC	2924
QY	1808	TTGACTTCAGGAGCTATGAATATGACTTAAAGTGGAGTTCCTGAGAGAGACTTTAGAGT	1867	Db	2906	CAGAAGAGCGATGATCAGAATGTGGATGGCGGTGTTTCGGAATGTCTCTCACACCTACC	2965
Db	1832	TTGATTTTCAAGAGATATGAATATGATCTCAATGGAGTTCCTGAGAGAGAAATTTAGAGT	1891	QY	2925	AAAAACAGGGCCCTCTAGCAGAGAGCGGGCTCAGAGC-CCAGTGGCCACAGCGCCAGG	2983
QY	1868	TTGGGAAGTCTCGGGTCTGGGCTTTTGGGAGGCTGATGAACGCCACCGCCTTATGGCA	1927	Db	2966	AAAAACAGGGCACCTTTTCCAGAGAGATGGATTTGGGGCTACTCTCTCCAGAGCTCAGG	3025
Db	1892	TTGGGAAGTCTAGATCAGTGTCTTTTGGAAAGTGTATGAACCAACAGCTTATGGAA	1951	QY	2984	T-CAAGATTTACAGAGAAAGAGTGTAGCGAGGAGGCTTTGGACCCCGCC-----ACCTAG	3038
QY	1928	TTAGTAAACCGGAGTCTCAATTTCAAGTGGCGGTGAAGATGCTTAAAGAGAAAGCTGACA	1987	Db	3026	TGAAGATTTCTAGAGGAACTTTAGTTTAAAGGACTTCTATCCCTCCACCTATCCCTAA	3085
Db	1952	TTAGCAAAACAGGAGTCTCAATCCAGGTTGCGCTCAAAATGCTGAAAGAAAGAGCAGACA	2011	QY	3039	CAGGCTGTAGACCGCAGAGCCAAAGATTAGCCTCGCCTCT--GAGGAAGCGCCCTACAGCG	3096
QY	1988	GCTGTGAAAAGAGCTCTCATGTGCGAGCTCAAAATGATGATGCCACCTGGGACACCATG	2047	Db	3086	CAGGCTGTAGATTACCAAAACAGATTAATTTTCACTACTTAAAGAAATCTATTATCAAC	3145
Db	2012	GCTCTGAAAAGAGAGCACTCATGTGAGAACTCAAGATGATGATGCCAGCTGGGAAGCCAG	2071	QY	3097	CGTTCCTTCTGCTGAGCTTTTCTCTAGATGCTGTGCTCAATCTC-----CAAGTGA	3149
QY	2048	ACAACATCGTGAATCTGCTGGGGGATGCACACTGTCTAGGGCCAGTGTACTTGTATTTTG	2107	Db	3146	TGCTGTCTTCAACAGACTTTTCTCTAGAAAGCCGCTCTGCGTTTACTTCTTGTGTTTCAAGGGA	3205
Db	2072	AGAAATATTTGAGACTGCTGTGGGGGGTGCACTGTCTAGGACCAATTTACTTGTATTTTG	2131	QY	3150	CTTCTATAAATAACAACTCTCTCTCGCACAGGCGGGAGAGCCAAATAATGAGACTTGTGG	3209
QY	2108	AATATTTGTTGTTGGTGAACCTCTCACTACCTAAGAGTAAAGAGAGAGAGTTTTCACA	2167				


```

; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: US 60/223,378
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: US 09/796,692
; PRIOR FILING DATE: 2001-03-01
; NUMBER OF SEQ ID NOS: 10467
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7538
; LENGTH: 645
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-040-862-7538

Query Match      13.5%; Score 456.6; DB 14; Length 645;
Best Local Similarity 84.5%; Pred. No. 5.1e-115; Mismatches 94; Indels 7; Gaps 2;
Matches 550; Conservative 0;

Qy 2035 CTGGACACCATGACCAACATCGTGAATCTGCTGGGGGATGACACACTGTGAGGGCCAGTG 2094
Db 1 CTGGGAAGCCACGAGATATTTGTGAACCTGCTGGGGGCTGACACTGTGAGGACCAATT 60

Qy 2095 TACTTGATTTTGAATATTTGTTGCTATGTTGACCTCTCAACTACTCTAAGAAAGTAAAGA 2154
Db 61 TACTTGATTTTGAATACTGTTGCTATGTTGATCTCTCAACTATCTAAGAAAGTAAAGA 120

Qy 2155 GAGAGTTTTCAGACACATGACAGATTTTTRAGGACATATTTTCAAGTCTTACCTT 2214
Db 121 GAAAAATTTTCAGAGACTTGGACAGAGATTTTCAAGGAAACACAAATTTACCTT 180

Qy 2215 ACTTTCCAGGACATTTCAAAATTCAGCATGCTGTTTCAAGGAAAGTTCAGTTACACCG 2274
Db 181 ACTTTCCAAATCACATCCAAATTCAGCATGCTGTTTCAAGGAAAGTTCAGTTACACCG 240

Qy 2275 CCCTTGGATCAGCTCTGAGGTTTCAATGGAAATTCATTTCACTCTGAAGATGAAATTGAA 2334
Db 241 GACTCGGATCAAAATCTCAGGGCTTCATGGGAATTCATTTCACTCTGAAGATGAAATTGAA 300

Qy 2335 TATGAAACCCAGAGAGCTGCGAGAGAGGAGGAGATTTTCAAGGAAAGTTCAGTTACACCG 2394
Db 301 TATGAAACCCAGAGAGCTGCGAGAGAGGAGGAGATTTTCAAGGAAAGTTCAGTTACACCG 354

Qy 2395 GAAGACCTCTCTTGTGTTTGGTTCGTTTCAAGGAAAGTTCAGTTACACCG 2454
Db 355 GAAGATCTCTTGTGTTTGGTTCGTTTCAAGGAAAGTTCAGTTACACCG 414

Qy 2455 TCGTGTGTTCCAGAGACCTGCGAGCCAGGAAATGTTGGTCCACCCAGGAGGTGGTG 2514
Db 415 TCGTGTGTTCCAGAGACCTGCGAGCCAGGAAATGTTGGTCCACCCAGGAGGTGGTG 474

Qy 2515 AAGATCTGTGACTTTGGACTGCGCCAGAGACATCTCTGAGGACTCCAGCTACGTCGAGG 2574
Db 475 AAGATATGTGACTTTGGATTTGGCTGCGAGATATCATGAGTATTCGAATCTGTTGTCAGG 534

Qy 2575 GGCACGACCGCTGCGCTGAGTGGATGGGACCCGAGAG-CTTATTTGAAGGATCTA 2633
Db 535 GGCATGCGCTGCTGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 594

Qy 2634 CACAAATCAAGAGTACGCTGCTGCTTCTGAGGATCTCTCTCTGAGGAGATTT 2684
Db 595 CACCAATTAAGAGTATGCTGCTGATATGGAATATTTACTGTGGGAATCTT 645

```

RESULT 8
 US-09-796-692-8085
 ; Sequence 8085, Application US/09796692
 ; Publication No. US20020198362A1
 ; GENERAL INFORMATION:

```

; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPI
; TITLE OF INVENTION: HEMATOLOGICAL MALIGNANCIES
; FILE REFERENCE: 2077.001200
; CURRENT APPLICATION NUMBER: US/09/796,692
; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8085
; LENGTH: 610
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-796-692-8085

Query Match      13.2%; Score 456.4; DB 10; Length 610;
Best Local Similarity 85.1%; Pred. No. 2.8e-112; Mismatches 86; Indels 6; Gaps 1;
Matches 524; Conservative 0;

Qy 2035 CTGGACACCATGACCAACATCGTGAATCTGCTGGGGGATGACACACTGTGAGGGCCAGTG 2094
Db 1 CTGGGAAGCCACGAGATATTTGTGAACCTGCTGGGGGCTGACACTGTGAGGACCAATT 60

Qy 2095 TACTTGATTTTGAATATTTGTTGCTATGTTGACCTCTCAACTACTCTAAGAAAGTAAAGA 2154
Db 61 TACTTGATTTTGAATACTGTTGCTATGTTGATCTCTCAACTATCTAAGAAAGTAAAGA 120

Qy 2155 GAGAAATTTTCAGGACATGACAGAGATTTTTRAGGAAACATAATTTCAAGTTCTTACCTT 2214
Db 121 GAAAAATTTTCAGAGACTTGGACAGAGATTTTCAAGGAAACACAAATTTCAAGTTCTTACCTT 180

Qy 2215 ACTTTCCAGGACATTTCAAAATTCAGCATGCTGTTTCAAGGAAAGTTCAGTTACACCG 2274
Db 181 ACTTTCCAAATCACATCCAAATTCAGCATGCTGTTTCAAGGAAAGTTCAGTTACACCG 240

Qy 2275 CCCTTGGATCAGCTCTGAGGTTTCAATGGAAATTCATTTCACTCTGAAGATGAAATTGAA 2334
Db 241 GACTCGGATCAAAATCTCAGGGCTTCATGGGAATTCATTTCACTCTGAAGATGAAATTGAA 300

Qy 2335 TATGAAACCCAGAGAGCTGCGAGAGAGGAGGAGATTTTCAAGGAAAGTTCAGTTACACCG 2394
Db 301 TATGAAACCCAGAGAGCTGCGAGAGAGGAGGAGATTTTCAAGGAAAGTTCAGTTACACCG 354

Qy 2395 GAAGACCTCTCTTGTGTTTGGTTCGTTTCAAGGAAAGTTCAGTTACACCG 2454
Db 355 GAAGATCTCTTGTGTTTGGTTCGTTTCAAGGAAAGTTCAGTTACACCG 414

Qy 2455 TCGTGTGTTCCAGAGACCTGCGAGCCAGGAAATGTTGGTCCACCCAGGAGGTGGTG 2514

```

Db 415 TCGTGTTCACAGAGACCTGGCCGAGCAAGTGTCTTGTACCCACGCGGAAGTGGT 474
QY 2515 AGATCTGTGACTTTGGACTGGCCGAGACATCTGAGCGACTCCAGCTACGTGTCAGG 2574
Db 475 AGATAATGTGACTTTGGATGGCTCGATATATCATGAGTGATTCACAACTATGTGTGAGG 534
QY 2575 GGCACCGCAGCGGTGCGCGGTGAAGTGGATGGCCACCCGAGAGCTTATTTGAAGGGATCTAC 2634
Db 535 GGCATGCGCGTGTGCTGTAAATGATGGCCGCCGGAAGCGTGTTTGAAGGCATCTAC 594
QY 2635 ACAATCAAGAGTGAGC 2650
Db 595 ACCATTAAAGATGATG 610

RESULT 9

US-10-040-862-8085
; Sequence 8085, Application US/10040862
; Publication No. US20030078396A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE REFERENCE: 014058-013520US
; CURRENT APPLICATION NUMBER: US/10/040,862
; CURRENT FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: US 60/223,378
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: US 09/796,692
; PRIOR FILING DATE: 2001-03-01
; NUMBER OF SEQ ID NOS: 10467
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8085
; LENGTH: 610
; TYPE: DNA
; ORGANISM: Homo sapiens

US-10-040-862-8085

Query Match 13.2%; Score 456.4; DB 14; Length 610;

Best Local Similarity 85.1%; Pred. No. 2.8e-112;

Matches 524; Conservative 0; Mismatches 86; Indels 6; Gaps 1;

QY 2035 CTGGGACCCATGACACATCTGCTGGGGGCGATGCACTGTGAGGGCCAGTG 2094

Db 1 CTGGGAGCCACGAGATATTTGAACTCTGCTGGGGGCGTGCACACTGTGAGGACCAATT 60

QY 2095 TACTGATTTTGAATATTGTTGCTATGTGACCTCTCTCACTACCTAAGATGAAAGA 2154
Db 61 TACTGATTTTGAATATTGTTGCTATGTGATCTTCTCACTATCTAAGATGAAAGA 120
QY 2155 GAGAACTTTTCAAGGACATGAGAGATTTTAAAGGACATATTTTCACTTCTTACCT 2214
Db 121 GAAATTTTCAAGGACTTTGGACAGAGATTTTCAAGGACATATTTTCACTTCTTACCT 180
QY 2215 ACTTTCCAGGACATTTCAAATTCAGCATGCTCTGTTTCAAGGAGATTTTCACTTCTTACCT 2274
Db 181 ACTTTCCAGGACATTTCAAATTCAGCATGCTCTGTTTCAAGGAGATTTTCACTTCTTACCT 240
QY 2275 CCTTTGGATCAGCTCTCAGGTTTCAATGGAAATTTCAATTTCTTCAAGATGAGATGAA 2334
Db 241 GACTGGATCAAATCTCAGGGCTTCTATGGGAATTTCACTTCTTCAAGATGAAATGAA 300
QY 2335 TATGAAACACAGAGAGGCTGCGAGAGAGAGGAGGAGATTTTGAACGTGCTGACGTTT 2394
Db 301 TATGAAACACAGAGAGGCTGCGAGAGAGAGGAGGAGATTTTGAACGTGCTGACGTTT 354
QY 2395 GAAGACCTCTTCTTGTCTTGTGCTACCAAGTGGCCAAAGGATGGAATTTCTGGAGTTCAAG 2454
Db 355 GAAGATCTTCTTGTCTTGTGCTATCAAGTTGCCAAGGATGGAATTTCTGGAGTTTAAAG 414
QY 2455 TCGTGTCTCCAGAGACCTGCGAGCCAGGAATGTTGGTCAACCACGGGAAAGTGGTG 2514
Db 415 TCGTGTCTCCAGAGACCTGCGAGCCAGGAATGTTGGTCAACCACGGGAAAGTGGTG 474
QY 2515 AGATCTGTGACTTTGGACTGGCCGAGACATCTTGGAGCATCTCAGCTACGTGTCAGG 2574
Db 475 AGATATGTGACTTTGGATTTGGCTCGATATATCATGAGTATTTCCAACACTATGTGTGAGG 534
QY 2575 GGCACCGCAGCGCTGCGGTGAAGTGGATGGCCACCCGAGAGCTTATTTCAAGGGATCTAC 2634
Db 535 GGCATGCGCGTGTGCTGTAAATGATGGCCGCCGGAAGCGTGTTTGAAGGCATCTAC 594
QY 2635 ACAATCAAGAGTGAGC 2650
Db 595 ACCATTAAAGATGATG 610

RESULT 10

US-09-796-692-8927
; Sequence 8927, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPEUTIC
; FILE REFERENCE: 2077.001200
; CURRENT APPLICATION NUMBER: US/09/796,692
; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903

;	TITLE OF INVENTION: Hematological Malignancies			
;	FILE REFERENCE: 014058-013520US			
;	CURRENT APPLICATION NUMBER: US/10/040,862			
;	CURRENT FILING DATE: 2001-11-06			
;	PRIOR APPLICATION NUMBER: US 60/186,126			
;	PRIOR FILING DATE: 2000-03-01			
;	PRIOR APPLICATION NUMBER: US 60/190,479			
;	PRIOR FILING DATE: 2000-03-17			
;	PRIOR APPLICATION NUMBER: US 60/200,545			
;	PRIOR FILING DATE: 2000-04-27			
;	PRIOR APPLICATION NUMBER: US 60/200,303			
;	PRIOR FILING DATE: 2000-04-28			
;	PRIOR APPLICATION NUMBER: US 60/200,779			
;	PRIOR FILING DATE: 2000-04-28			
;	PRIOR APPLICATION NUMBER: US 60/200,999			
;	PRIOR FILING DATE: 2000-05-01			
;	PRIOR APPLICATION NUMBER: US 60/202,084			
;	PRIOR FILING DATE: 2000-05-04			
;	PRIOR APPLICATION NUMBER: US 60/206,201			
;	PRIOR FILING DATE: 2000-05-22			
;	PRIOR APPLICATION NUMBER: US 60/218,950			
;	PRIOR FILING DATE: 2000-07-14			
;	PRIOR APPLICATION NUMBER: US 60/222,903			
;	PRIOR FILING DATE: 2000-08-03			
;	PRIOR APPLICATION NUMBER: US 60/223,416			
;	PRIOR FILING DATE: 2000-08-04			
;	PRIOR APPLICATION NUMBER: US 60/223,378			
;	PRIOR FILING DATE: 2000-08-07			
;	PRIOR APPLICATION NUMBER: US 09/796,692			
;	PRIOR FILING DATE: 2001-03-01			
;	NUMBER OF SEQ ID NOS: 10467			
;	SOFTWARE: FastSeq for Windows Version 3.0			
;	SEQ ID NO 8927			
;	LENGTH: 605			
;	TYPE: DNA			
;	ORGANISM: Homo sapiens			
;	US-10-040-862-8927			
Query Match	13.2%	Score 454.6;	DB	
Best Local Similarity	85.3%	Pred. No. 8.4e-11		
Matches 521;	Conservative	0;	Mismatches	
Qy	2035	CTGGGACACCATGACAAACATCGTGAATCTCTCTGGGGG		
Db	1	CTGGGAAGCCACGAGATATTGTGAACCTCTCTGGGGG		
Qy	2095	TACTTGATTTTTGCAATATCTTGTCTATGGTGACCTCC		
Db	61	TACTTGATTTTTGAACTACTTGTCTATGGTGATCTTC		
Qy	2155	GAGAAGTTTTCACAGGACATGGACAGAGATTTTTTAAGG		
Db	121	GAAAATTTTCACAGGACTTGGACAGAGATTTTCAAGG		
Qy	2215	ACTTTCAGGCAATTCAAATTCACGATGCCTGGTTT		
Db	181	ACTTTCGAATCATCCAAATTCACGATGCCTGGTTT		
Qy	2275	CCCTTGGATCGCTCTCAGGGTTCAATGGGAATTCAA		
Db	241	GATCTGGATCAAAATCTCAGGGTTTCATGGGAATTCAT		
Qy	2335	TATGAAAACACAGAAGGCTGGCAGAAGAAGAGGAGG		
Db	301	TATGAAAACCAAAAAGGCTG-----GAAGAAGAGG		
Qy	2395	GAAGACCTCTTTTGGCTTTTCGTCACCAAGTGGCCAAAG		
Db	355	GAAGATCTCTTTTGGCTTTTGCATATCAAGTTGCCAAAG		
Qy	2455	TCGTGTGTCCACAGAGACCTGGCAGCCAGGAATGTGT		
Db	415	TCGTGTGTTCACAGAGACCTGGCCGACGAGAACGTGC		


```
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: US 09/796,692
; PRIOR FILING DATE: 2001-03-01
; NUMBER OF SEQ ID NOS: 10467
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7976
; LENGTH: 597
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-040-862-7976

Query Match      13.0%; Score 448.8; DB 14; Length 597;
Best Local Similarity 85.4%; Pred. No. 3.1e-110;
Matches 514; Conservative 0; Mismatches 82; Indels 6; Gaps 1;

QY 2035 CTGGGACACCATGACAACTCGTGAATCTCTGGGGCATGCACACTGTCCAGGCCAGTG 2094
Db 1 CTGGGAGCCACGAGAAATATGTGAACCTCTCTGGGGCGTGCACACTGTCCAGGACCAATT 60

QY 2095 TACTTGATTTTGAATATTTGCTATGTGTGACCTCTCAACTACCTTAAGAACTAAAGA 2154
Db 61 TACTTGATTTTGAATATTTGCTATGTGTGATCTTCTCAACTATCTAAGAACTAAAGA 120

QY 2155 GAGAAGTTTTCACAGGACATGGACAGAGATTTTAAAGGAACATAATTTCAAGTTTCAAG 2454
Db 121 GAAAAATTTTCACAGGACTTGGACAGAGATTTTCAAGGAACACAAATTTCAAGTTTCAAG 180

QY 2215 ACTTTCCAGGCACATTCAAATTCAGCATGCTGCTTCAAGCAAGTTCAGTTACACCG 2274
Db 181 ACTTTCCAAATCACATCCAAATTCAGCATGCTGCTTCAAGCAAGTTCAGTTACACCG 240

QY 2275 CCCTTGGATCAGCTCTCAGGGTTCAATCGGAATTCATTTCTGAAGATGAGTTGAA 2334
Db 241 GACTCGGATCAATCTCAGGGCTTCATGGGAATTCATTTCACTCTGAAGATGAAATTGAA 300

QY 2335 TATGAAACCCAGAGAGCTGGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2394
Db 301 TATGAAACCCAGAGAGAGCTGGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 354

QY 2395 GAAGACCTCTCTTCTGCTTTCGATACCAAGTGGCCAAAGGATGGAATTCCTGGAGTTCAAG 2454
Db 355 GAAGATCTCTTCTGCTTTCGATATCAAGTTTCCAAAGGAATGGAATTTCTGGAATTTAAG 414

QY 2455 TCGTGTGTTCACAGACCTGGCAGCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2514
Db 415 TCGTGTGTTCACAGACCTGGCAGCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 474

QY 2515 AAGATCTGTGACTTTGGACTTGGCCCGGAGACATCTCTGAGCGAGACTCCAGTACCTGCTCAGG 2574
Db 595 AC 596

QY 2575 GGCACCGCAGCGCTCCCGTGAAGTGGATGGCACCAGAGAGTTATTTGAAGGAGATCTAC 2634
Db 535 GGCATGCGCGCTCTGCTGTAAATGGATGGCCCGGAGAGAGAGAGAGAGAGAGAGAGAGAG 594

QY 2635 AC 2636

Db 595 AC 596
```

RESULT 14

```
US-09-796-692-8273
; Sequence 8273; Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
; FILE REFERENCE: 2077.001200
; CURRENT APPLICATION NUMBER: US/09/796,692
; CURRENT FILING DATE: 2001-03-01
```

```
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8273
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (586)
; OTHER INFORMATION: n=A,T,C or G
US-09-796-692-8273
```

```
Query Match      13.0%; Score 448; DB 10; Length 601;
Best Local Similarity 84.8%; Pred. No. 5.1e-110;
Matches 515; Conservative 0; Mismatches 86; Indels 6; Gaps 1;

QY 2035 CTGGGACACCATGACAACTCGTGAATCTCTGGGGCATGCACACTGTCCAGGCCAGTG 2094
Db 1 CTGGGAGCCACGAGAAATATGTGAACCTCTCTGGGGCGTGCACACTGTCCAGGACCAATT 60

QY 2095 TACTTGATTTTGAATATTTGCTATGTGTGACCTCTCAACTACCTTAAGAACTAAAGA 2154
Db 61 TACTTGATTTTGAATATTTGCTATGTGTGATCTTCTCAACTATCTAAGAACTAAAGA 120

QY 2155 GAGAAGTTTTCACAGGACATGGACAGAGATTTTAAAGGAACATAATTTCAAGTTTCAAG 2214
Db 121 GAAAAATTTTCACAGGACTTGGACAGAGATTTTCAAGGAACACAAATTTCAAGTTTCAAG 180

QY 2215 ACTTTCCAGGCACATTCAAATTCAGCATGCTGCTTCAAGCAAGTTCAGTTACACCG 2274
Db 181 ACTTTCCAAATCACATCCAAATTCAGCATGCTGCTTCAAGCAAGTTCAGTTACACCG 240

QY 2275 CCCTTGGATCAGCTCTCAGGGTTCAATCGGAATTCATTTCTGAAGATGAGTTGAA 2334
Db 241 GACTCGGATCAATCTCAGGGCTTCATGGGAATTCATTTCACTCTGAAGATGAAATTGAA 300

QY 2335 TATGAAACCCAGAGAGCTGGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2394
Db 301 TATGAAACCCAGAGAGAGCTGGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 354

QY 2395 GAAGACCTCTCTTCTGCTTTCGATACCAAGTGGCCAAAGGATGGAATTCCTGGAGTTCAAG 2454
Db 355 GAAGATCTCTTCTGCTTTCGATATCAAGTTTCCAAAGGAATGGAATTTCTGGAATTTAAG 414

QY 2455 TCGTGTGTTCACAGACCTGGCAGCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2514
Db 415 TCGTGTGTTCACAGACCTGGCAGCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 474

QY 2515 AAGATCTGTGACTTTGGACTTGGCCCGGAGACATCTCTGAGCGAGACTCCAGTACCTGCTCAGG 2574
```


Db 475 AAGATATGACCTTTGGATTGCTCGATATATCATGAGTATCCAACTATGTTGTCAGG 534
 Qy 2575 GGCAACGACCGCTCCGGTGAAGTGGATGGCACCCGAGAGCTTATTTGAAGGGATCTAC 2634
 Db 535 GGCAATGCCGCTGCTGCTGTAATAAGTGGTGGCCCCGGAAGCCTGTTTGAANGCATCTAC 594
 Qy 2635 ACAATCA 2641
 Db 595 ACCATTA 601

RESULT 15
 US-10-040-862-8273
 ; Sequence 8273, Application US/10040862
 ; Publication No. US20030078396A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Gaiger, Alexander
 ; APPLICANT: Algate, Paul A.
 ; APPLICANT: Mannion, Jane
 ; APPLICANT: Retter, Marc
 ; APPLICANT: Corixa Corporation
 ; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
 ; OF INVENTION: Hematological Malignancies
 ; FILE REFERENCE: 014058-013520US
 ; CURRENT APPLICATION NUMBER: US/10/040,862
 ; CURRENT FILING DATE: 2001-11-06
 ; PRIOR APPLICATION NUMBER: US 60/186,126
 ; PRIOR FILING DATE: 2000-03-01
 ; PRIOR APPLICATION NUMBER: US 60/190,479
 ; PRIOR FILING DATE: 2000-03-17
 ; PRIOR APPLICATION NUMBER: US 60/200,545
 ; PRIOR FILING DATE: 2000-04-27
 ; PRIOR APPLICATION NUMBER: US 60/200,303
 ; PRIOR FILING DATE: 2000-04-28
 ; PRIOR APPLICATION NUMBER: US 60/200,779
 ; PRIOR FILING DATE: 2000-04-28
 ; PRIOR APPLICATION NUMBER: US 60/200,999
 ; PRIOR FILING DATE: 2000-05-01
 ; PRIOR APPLICATION NUMBER: US 60/202,084
 ; PRIOR FILING DATE: 2000-05-04
 ; PRIOR APPLICATION NUMBER: US 60/206,201
 ; PRIOR FILING DATE: 2000-05-22
 ; PRIOR APPLICATION NUMBER: US 60/218,950
 ; PRIOR FILING DATE: 2000-07-14
 ; PRIOR APPLICATION NUMBER: US 60/222,903
 ; PRIOR FILING DATE: 2000-08-03
 ; PRIOR APPLICATION NUMBER: US 60/223,416
 ; PRIOR FILING DATE: 2000-08-04
 ; PRIOR APPLICATION NUMBER: US 60/223,378
 ; PRIOR FILING DATE: 2000-08-07
 ; PRIOR APPLICATION NUMBER: US 09/796,692
 ; PRIOR FILING DATE: 2001-03-01
 ; NUMBER OF SEQ ID NOS: 10467
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 8273
 ; LENGTH: 601
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: unsure
 ; LOCATION: (586)
 ; OTHER INFORMATION: n=A,T,C or G
 US-10-040-862-8273

Query Match 13.0%; Score 448; DB 14; Length 601;
 Best Local Similarity 84.8%; Pred. No. 5.1e-110;
 Matches 515; Conservative 0; Mismatches 86; Indels 6; Gaps 1;
 Qy 2035 CTGGACACCATGACAACTGTAATCTGTGGGGGATGACACTGTGACGGCCAGTG 2094
 Db 1 CTGGGAAGCCAGAGATATTTGGAACCTGCTGGGGGGCTGTGACACTGTGACGACCAAT 60

Qy 2095 TACTTTGATTTTGAATATTTGTTGCTATGTTGACCTCTCAACTACCTAAGAAAGTAAAGA 2154
 Db 61 TACTTTGATTTTGAATATTTGTTGCTATGTTGACCTCTCAACTATCTAAGAAAGTAAAGA 120
 Qy 2155 GAGAAGTTTTCACAGGACATGGACAGAGATTTTAAAGGAACAATAATTTCAAGTTTCAACCT 2214
 Db 121 GAAAAATTTTCACAGGACATTTGGACAGAGATTTTCAAGGAACAACAATTTCAAGTTTCAACCT 180
 Qy 2215 ACTTTCCAGGACATTTCAAAATTCAGCATGCTGTTTCAAGGAAGTTCAGTTTACACCTG 2274
 Db 181 ACTTTCCAAATCACATCCAAATTTCCAGCATGCTGTTTCAAGGAAGTTCAGATACACCTG 240
 Qy 2275 CCTTTGGATCAGCTCTCAGGGTTCAATGGGAATTTCAATTTCTTGAAGATGAGATTGAA 2334
 Db 241 GACTCGGATCAAAATCTCAGGGCTTCATGGGAATTTCAATTTCACTCTGGAAGATGAAATTGAA 300
 Qy 2335 TATGAAAAACCAAGAGAGCTGGCAGAGAAGAAAGAGGAGGAATTTGAAACGCTGCTGACGTTT 2394
 Db 301 TATGAAAAACCAAAAGGCTG-----GAAGAGAGGAGGACTTTGAATGCTGCTTACATTT 354
 Qy 2395 GAAGACCTCTTTGCTTTGCTTACCAAGTGGCCAAAGGAGGATTTCTGAGATTCAAG 2454
 Db 355 GAAGATCTTTCTTTGCTTATCAAGTTGCCAAAGGAATGGAATTTCTGGAATTTAAG 414
 Qy 2455 TCGTGTGTCACAGAGACCTGGCAGCAGGATGTTGGTCCACCCACGGAAGTGGTG 2514
 Db 415 TCGTGTGTTTACAGAGACCTGGCCCGCCAGGAACGTTGTTCACCCACGGAAGTGGTG 474
 Qy 2515 AAGATCTGTGACTTTGGGACTGGCCCGGAGACATCTGAGCGACTCCAGCTACGTCGTCAGG 2574
 Db 475 AAGATATGTGACTTTGGATTTGGCTCGATATATCATGAGTGTATTTCCAACTATGTTGTCAGG 534
 Qy 2575 GGCAACGACCGCTCCGGTGAAGTGGATGGCACCCGAGAGCTTATTTGAAGGGATCTAC 2634
 Db 535 GGCAATGCCGCTGCTGCTGTAATAAGTGGATGGCCCGGAAAGCCTGTTTGAANGCATCTAC 594
 Qy 2635 ACAATCA 2641
 Db 595 ACCATTA 601

Search completed: August 28, 2003, 13:26:34
 Job time : 1016.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 27, 2003, 15:21:07 ; Search time 280.55 Seconds

(without alignments)
5432.528 Million cell updates/sec

Title: US-09-919-408A-1

Perfect score: 3453

Sequence: 1 GCGGCTGCTACCGCGCGC.....AAAAAAAAAAAAAAAAAAAA 3453

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 569378 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA:*

1: /cgn2_6/prodata/2/ina/5A COMB.seq.*

2: /cgn2_6/prodata/2/ina/5B COMB.seq.*

3: /cgn2_6/prodata/2/ina/6A COMB.seq.*

4: /cgn2_6/prodata/2/ina/6B COMB.seq.*

5: /cgn2_6/prodata/2/ina/6C COMB.seq.*

6: /cgn2_6/prodata/2/ina/6D COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3453	100.0	3453	1	US-07-813-593-1
2	3453	100.0	3453	1	US-07-977-451-1
3	3453	100.0	3453	1	US-07-946-507-1
4	3453	100.0	3453	1	US-08-252-517-1
5	3453	100.0	3453	1	US-07-906-397A-1
6	3453	100.0	3453	1	US-08-601-891-1
7	3453	100.0	3453	2	US-09-021-324-1
8	3453	100.0	3453	5	PCT-US92-02750-1
9	3453	100.0	3453	5	PCT-US92-05401-1
10	3453	100.0	3453	5	PCT-US92-09893-1
11	3344	96.8	3521	1	US-08-222-299-1
12	3344	96.8	3521	2	US-08-434-878-1
13	3344	96.8	3521	5	PCT-US95-03718-1
14	3227.2	64.5	3501	1	US-07-977-451-3
15	2227.2	64.5	3501	1	US-08-252-517-3
16	2227.2	64.5	3501	1	US-07-906-397A-3
17	2227.2	64.5	3501	1	US-08-601-891-3
18	2227.2	64.5	3501	2	US-09-021-324-3
19	2227.2	64.5	3501	5	PCT-US92-05401-3
20	2227.2	64.5	3501	5	PCT-US92-09893-3
21	2204.8	63.9	3476	1	US-08-183-211-1
22	2204.8	63.9	3476	5	PCT-US95-00176A-1
23	2203	63.8	3475	1	US-08-222-299-3
24	2203	63.8	3475	2	US-08-434-878-3
25	2203	63.8	3475	5	PCT-US95-03718-3
26	2081.8	60.3	3120	1	US-08-222-616-22
27	2081.8	60.3	3120	4	US-08-446-648-22

28 2081.8 60.3 3120 5 PCT-US95-04228-22 Sequence 22, Appli

29 1206 34.9 1894 3 US-07-912-122-3 Sequence 3, Appli

30 1206 34.9 1894 5 PCT-US93-06404-3 Sequence 3, Appli

31 329 9.5 5084 1 US-08-306-691B-21 Sequence 21, Appli

32 329 9.5 5084 5 PCT-US93-06251-25 Sequence 25, Appli

33 286.4 8.3 4054 1 US-08-180-195-35 Sequence 35, Appli

34 286.4 8.3 4054 1 US-08-477-329-35 Sequence 35, Appli

35 286.4 8.3 4054 2 US-08-475-458-35 Sequence 35, Appli

36 286.4 8.3 4054 3 US-08-980-400-35 Sequence 35, Appli

37 286.4 8.3 4054 3 US-09-583-459A-35 Sequence 35, Appli

38 286.4 8.3 4054 3 US-09-583-210-35 Sequence 35, Appli

39 286.4 8.3 4054 4 US-09-583-449A-35 Sequence 35, Appli

40 286.4 8.3 4054 4 US-09-435-059-35 Sequence 35, Appli

41 286.4 8.3 4100 1 US-08-168-917-3 Sequence 3, Appli

42 286.4 8.3 4100 2 US-08-460-510-3 Sequence 3, Appli

43 286.4 8.3 4100 2 US-08-460-490-3 Sequence 3, Appli

44 286.4 8.3 4100 5 PCT-US92-00730-3 Sequence 3, Appli

45 286.4 8.3 4100 5 PCT-US92-00862-3 Sequence 3, Appli

ALIGNMENTS

RESULT 1

US-07-813-593-1

; Sequence 1, Application US/07813593

; Patent No. 5185438

; GENERAL INFORMATION:

; APPLICANT: Lemischka, Ihor R.

; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL

; TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS

; NUMBER OF SEQUENCES: 4

; CORRESPONDENCE ADDRESS:

; ADDRESSER: IMCLONE SYSTEMS INCORPORATED

; STREET: 180 VARICK STREET

; CITY: NEW YORK

; STATE: NEW YORK

; COUNTRY: U.S.A.

; ZIP: 10014

; COMPUTER READABLE FORM: disk

; MEDIUM TYPE: Floppy

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/07/813,593

; FILING DATE: 19920415

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/793,065

; FILING DATE: 15-NOV-1991

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/728,913

; FILING DATE: 28-JUN-1991

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/679,666

; FILING DATE: 02-APR-1991

; ATTORNEY/AGENT INFORMATION:

; NAME: Feit, Irving N.

; REGISTRATION NUMBER: 28,601

; REFERENCE/DOCKET NUMBER: LEM-3-PPP

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 212-645-1405

; TELEFAX: 212-645-2054

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 3453 base pairs

; TYPE: NUCLEIC ACID

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: cDNA

; FEATURE:

; NAME/KEY: CDS

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/679,666
FILING DATE: 02-APR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Feit, Irving N.
REGISTRATION NUMBER: 28,601
REFERENCE/DOCKET NUMBER: LEM-3-7P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-645-1405
TELEFAX: 212-645-2054
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3453 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 112..3006
NAME/KEY: sig_peptide
LOCATION: 31..111
FEATURE:
NAME/KEY: CDS
LOCATION: 31..3009
US-07-977-451-1

Query Match 100.0%; Score 3453; DB 1; Length 3453;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3453; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCGGCTGTGCTACCGCGGCTCCGAGGCGATCGCGGCGTTGGCGCAGCGCAGCGACCGG 60
DB 1 GCGGCTGTGCTACCGCGGCTCCGAGGCGATCGCGGCGTTGGCGCAGCGCAGCGACCGG 60
QY 61 GCGGCTGTGCTGTGTTGTTGTCAGTAATGATTTGAGACCGTTTACAAACCAAGAC 120
DB 61 GCGGCTGTGCTGTGTTGTTGTCAGTAATGATTTGAGACCGTTTACAAACCAAGAC 120
QY 121 CTGCTGTGATCAAGTGTGTTTATTCAGTCAATGAGAAATGCGTCTATCAGCGGGAAG 180
DB 121 CTGCTGTGATCAAGTGTGTTTATTCAGTCAATGAGAAATGCGTCTATCAGCGGGAAG 180
QY 181 CCATCATCGTACCGAATGTCGAGGATCCCGAAGACCTCCAGTGTACCCCGAGGCGC 240
DB 181 CCATCATCGTACCGAATGTCGAGGATCCCGAAGACCTCCAGTGTACCCCGAGGCGC 240
QY 241 CAGAGTGAAGGACCGGTATATGAAGCGGCACCGTGGAGGTGGCGAGTCTGGGTCCATC 300
DB 241 CAGAGTGAAGGACCGGTATATGAAGCGGCACCGTGGAGGTGGCGAGTCTGGGTCCATC 300
QY 301 ACCCTGCAAGTGCAGCTGCCACCCGAGGACCTTCTGCTGTGGTCTTTAAGCAC 360
DB 301 ACCCTGCAAGTGCAGCTGCCACCCGAGGACCTTCTGCTGTGGTCTTTAAGCAC 360
QY 361 AGCTCCCTGGGCTGCGACCGCACCTTTGATTTACAAAACAGAGGAATCGTTTCCATGGCC 420
DB 361 AGCTCCCTGGGCTGCGACCGCACCTTTGATTTACAAAACAGAGGAATCGTTTCCATGGCC 420
QY 421 ATCTTGAAACGTGACAGACCCAGGACGAGTAATACCTACTTATTCAGAGCGAAGCGC 480
DB 421 ATCTTGAAACGTGACAGACCCAGGACGAGTAATACCTACTTATTCAGAGCGAAGCGC 480
QY 481 GCCAACTACAGTACTGTTTACAGTAATGTAAGATACACAGTGTATGCTAAGG 540
DB 481 GCCAACTACAGTACTGTTTACAGTAATGTAAGATACACAGTGTATGCTAAGG 540
QY 541 AGACCTTACTTTAGGAAGATGGAAACAGGATGACCTGCTGTCATCTCCGAGGCTTT 600
DB 541 AGACCTTACTTTAGGAAGATGGAAACAGGATGACCTGCTGTCATCTCCGAGGCTTT 600

DB 541 AGACCTTACTTTAGGAAGATGGAAACAGGATGACCTGCTGTCATCTCCGAGGCTTT 600
QY 601 CCGAGAGCCCACTGTGAGTGGGTGCTCTCAGCTCCCAAGGAAAGCTGTGTAAGAAGAA 660
DB 601 CCGAGAGCCCACTGTGAGTGGGTGCTCTCAGCTCCCAAGGAAAGCTGTGTAAGAAGAA 660
QY 661 GCGCCCTGCTGTGTCAGAAAGGAGGAAAGGTACTTTCATGAGTTGTCGGAACAGACATC 720
DB 661 GCGCCCTGCTGTGTCAGAAAGGAGGAAAGGTACTTTCATGAGTTGTCGGAACAGACATC 720
QY 721 AGATGCTGTGTAGAAATGCACTGGGCCCGCAATGCCAAGCTGTTCCACATAGATCTA 780
DB 721 AGATGCTGTGTAGAAATGCACTGGGCCCGCAATGCCAAGCTGTTCCACATAGATCTA 780
QY 781 AACAGGCTCTCAGAGCACACTGCCCAAGTTATTCCTGAAAGTGGGGAAACCTTTGTGG 840
DB 781 AACAGGCTCTCAGAGCACACTGCCCAAGTTATTCCTGAAAGTGGGGAAACCTTTGTGG 840
QY 841 ATCAGGTGTAAGGCCATCCATGTGAACCATGATTCGGGCTCACCTGGGAGCTGGAAGAC 900
DB 841 ATCAGGTGTAAGGCCATCCATGTGAACCATGATTCGGGCTCACCTGGGAGCTGGAAGAC 900
QY 901 AAAGCCCTGGAGGAGGCGAGCTACTTTGAGATGAGTACCTCTCCACAAACAGACCATG 960
DB 901 AAAGCCCTGGAGGAGGCGAGCTACTTTGAGATGAGTACCTCTCCACAAACAGACCATG 960
QY 961 ATTCCGATTTCTTGGGCTTTGTGTTCTCCGTTGGGAGAAACGACACCGGATATTACACC 1020
DB 961 ATTCCGATTTCTTGGGCTTTGTGTTCTCCGTTGGGAGAAACGACACCGGATATTACACC 1020
QY 1021 TGCTCTTCTCAAGACACCCAGGCGAGTCAGGCTTGGTGACCATCTCTAGAAAAGGTTT 1080
DB 1021 TGCTCTTCTCAAGACACCCAGGCGAGTCAGGCTTGGTGACCATCTCTAGAAAAGGTTT 1080
QY 1081 ATAAACGGTACCAGCTCGCAAGAGATGATAAATTGACCGTACGAAAGTTCTGCTTC 1140
DB 1081 ATAAACGGTACCAGCTCGCAAGAGATGATAAATTGACCGTACGAAAGTTCTGCTTC 1140
QY 1141 TCAGTCAGGTTTAAAGCGTACCCACGAATCCGATGCGATGCTTCTCTCAAGCCTCA 1200
DB 1141 TCAGTCAGGTTTAAAGCGTACCCACGAATCCGATGCGATGCTTCTCTCAAGCCTCA 1200
QY 1201 TTTCTTGTGAACAGAGGCTGAGGATGGGTACAGCATATCTAAATTTTTCGATCAT 1260
DB 1201 TTTCTTGTGAACAGAGGCTGAGGATGGGTACAGCATATCTAAATTTTTCGATCAT 1260
QY 1261 AAGAACAGCCAGGAGAGTACATATTTCTATGAGAAATGATGACGCCAGTTTCCACAA 1320
DB 1261 AAGAACAGCCAGGAGAGTACATATTTCTATGAGAAATGATGACGCCAGTTTCCACAA 1320
QY 1321 ATGTTTCAAGCTGAATATAGAAAGAAACCTCAAGTGTAGCAAAATGCTCAGCCAGCCAG 1380
DB 1321 ATGTTTCAAGCTGAATATAGAAAGAAACCTCAAGTGTAGCAAAATGCTCAGCCAGCCAG 1380
QY 1381 GGTCTCTGTTCTCTGATGGCTACCCGCTACCCCTCTTGGACCTGGAAGAGTTTCGGAC 1440
DB 1381 GGTCTCTGTTCTCTGATGGCTACCCGCTACCCCTCTTGGACCTGGAAGAGTTTCGGAC 1440
QY 1441 AAATCTCCCAATTCACCGAGGAAATCCAGAGAGGTTTGGAAATAAAGGCTAACAGA 1500
DB 1441 AAATCTCCCAATTCACCGAGGAAATCCAGAGAGGTTTGGAAATAAAGGCTAACAGA 1500
QY 1501 AAAGTGTGTCAGTGGCTGTCAGAGCAGTACTCTAAATATGAGTGAGGCGGGGAAAGG 1560
DB 1501 AAAGTGTGTCAGTGGCTGTCAGAGCAGTACTCTAAATATGAGTGAGGCGGGGAAAGG 1560
QY 1561 CTTCTGCTCAATGCTGTGCTACAAATTTCTATGGGCAAGTCTTGCAGAACCACTTTTTTA 1620
DB 1561 CTTCTGCTCAATGCTGTGCTACAAATTTCTATGGGCAAGTCTTGCAGAACCACTTTTTTA 1620
QY 1621 AACTCAGCAGGCGCTTCCCTTTCTATCAAGACACATCTCTCTATGCGGACCATGGG 1680
DB 1621 AACTCAGCAGGCGCTTCCCTTTCTATCAAGACACATCTCTCTATGCGGACCATGGG 1680

Thu Aug 28 13:01:19 2003

```

QY 1681 CTCTGTCTCCCTTCATTTGTTTCTCATTTGTTGTTGATCTGCCA CAAATACAAAAGCAA 1740
Db 1681 CTCTGTCTCCCTTCATTTGTTTCTCATTTGTTGTTGATCTGCCA CAAATACAAAAGCAA 1740
QY 1741 TTTAGTACGAGAGTCAGTCGAGATGATCAGAGTACCTGGCCCTCGATACCGATAC 1800
Db 1741 TTTAGTACGAGAGTCAGTCGAGATGATCAGAGTACCTGGCCCTCGATACCGATAC 1800
QY 1801 TTCTACGTTGACTTCCAGGAGCTATGATATGATACCTTAAGTGGAGTTCCTCGAGAGAGAAC 1860
Db 1801 TTCTACGTTGACTTCCAGGAGCTATGATATGATACCTTAAGTGGAGTTCCTCGAGAGAGAAC 1860
QY 1861 TTAGAGTTTGGAGAGTCTGGCGTCTGGCGTCTGGCGTCTGGCGTCTGGCGTCTGGCGTCTGGCG 1920
Db 1861 TTAGAGTTTGGAGAGTCTGGCGTCTGGCGTCTGGCGTCTGGCGTCTGGCGTCTGGCGTCTGGCG 1920
QY 1921 TATGCACTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1980
Db 1921 TATGCACTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1980
QY 1981 GCTGACAGCTGTAAGGAGTCTCATGTCGGAGCTCAAAATGATGATGATGATGATGATGATGATGATGATGAT 2040
Db 1981 GCTGACAGCTGTAAGGAGTCTCATGTCGGAGCTCAAAATGATGATGATGATGATGATGATGATGATGATGAT 2040
QY 2041 CACCATGACAACTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2100
Db 2041 CACCATGACAACTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2100
QY 2101 ATTTTGAATATGTTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2160
Db 2101 ATTTTGAATATGTTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2160
QY 2161 TTTACAGACATGACAGAGATTTTAAAGACATATTTTACAGTCTTACCTTACCTTACCTTACCTTACCT 2220
Db 2161 TTTACAGACATGACAGAGATTTTAAAGACATATTTTACAGTCTTACCTTACCTTACCTTACCTTACCT 2220
QY 2221 CAGGACATTTCAAAATCCAGCATGCTGTTTACAGAGAGTTCAGTTTACAGTTCAGTTTACAGTTCAGTT 2280
Db 2221 CAGGACATTTCAAAATCCAGCATGCTGTTTACAGAGAGTTCAGTTTACAGTTCAGTTTACAGTTCAGTT 2280
QY 2281 GATCAGCTCTCAGGTTCAATGGAATTCATTTCTGAGATGATGATGATGATGATGATGATGATGATGATGAT 2340
Db 2281 GATCAGCTCTCAGGTTCAATGGAATTCATTTCTGAGATGATGATGATGATGATGATGATGATGATGATGAT 2340
QY 2341 AACCAAGAGGCTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2400
Db 2341 AACCAAGAGGCTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2400
QY 2401 CTCCTTTGCTTTCGTTACCAAGTGGCCAAAGGAGTGAATTTCTTGGAGTTTCAAGTCTGTT 2460
Db 2401 CTCCTTTGCTTTCGTTACCAAGTGGCCAAAGGAGTGAATTTCTTGGAGTTTCAAGTCTGTT 2460
QY 2461 GTCCACAGAGACCTGGCAGCAGGAGATGTTGTTGTTTACCCACCCAGGAGGTTGTTGAGATC 2520
Db 2461 GTCCACAGAGACCTGGCAGCAGGAGATGTTGTTGTTTACCCACCCAGGAGGTTGTTGAGATC 2520
QY 2521 TGTGACTTTGGACTGGCCCGAGAGATCCTGAGGAGTCCAGTACCTGAGTACCTGAGTACCTGAGTACCTGAG 2580
Db 2521 TGTGACTTTGGACTGGCCCGAGAGATCCTGAGGAGTCCAGTACCTGAGTACCTGAGTACCTGAGTACCTGAG 2580
QY 2581 GCACGGCTGCGGTGAGTGGATGGATGGATGGATGGATGGATGGATGGATGGATGGATGGATGGATGGATGGAT 2640
Db 2581 GCACGGCTGCGGTGAGTGGATGGATGGATGGATGGATGGATGGATGGATGGATGGATGGATGGATGGATGGAT 2640
QY 2641 AAGAGTACAGTCTGCTGCTACCGCATCTTCTCTGGGAGATATTTTCACTGGGTGTAAC 2700
Db 2641 AAGAGTACAGTCTGCTGCTACCGCATCTTCTCTGGGAGATATTTTCACTGGGTGTAAC 2700
QY 2701 CTTTACCTGTCATCTCTGTCGAGCGTAACTTCTTATAACTGATTCAGAGTGGATTTTAA 2760
Db 2701 CTTTACCTGTCATCTCTGTCGAGCGTAACTTCTTATAACTGATTCAGAGTGGATTTTAA 2760

```

```

QY 2761 ATGGAGCAGCCATTTCTATGCCACAGAGGGATATATTTTGTAAATGCAATCTCTGCTGGGCT 2820
Db 2761 ATGGAGCAGCCATTTCTATGCCACAGAGGGATATATTTTGTAAATGCAATCTCTGCTGGGCT 2820
QY 2821 TTTTGTCTCAAGAGAGGGGCTATCTTCCCAACCTGACTTCAATTTTGTAGATGTCAGCTG 2880
Db 2821 TTTTGTCTCAAGAGAGGGGCTATCTTCCCAACCTGACTTCAATTTTGTAGATGTCAGCTG 2880
QY 2881 GCAGAGGCGAGAGAGAGTATCAGAACATCATCTACCAAAACAGGCGGCGCCCT 2940
Db 2881 GCAGAGGCGAGAGAGAGTATCAGAACATCATCTACCAAAACAGGCGGCGCCCT 2940
QY 2941 CAGCAGAGAGGGGCTCAGAGCCCGAGTCCGACAGCCAGGTCAGAGTGAAGTTCACAGAGAA 3000
Db 2941 CAGCAGAGAGGGGCTCAGAGCCCGAGTCCGACAGCCAGGTCAGAGTGAAGTTCACAGAGAA 3000
QY 3001 AGAAGTTAGCAGAGGAGGCTTGGACCCCGCCAGTCCAGAGGCTGTAGACCGCAGAGCCA 3060
Db 3001 AGAAGTTAGCAGAGGAGGCTTGGACCCCGCCAGTCCAGAGGCTGTAGACCGCAGAGCCA 3060
QY 3061 AGATTAGCTCGCTCTGAGGAGGCGCTTACAGCGGCTTCTTCTGCTGAGCTTTTCTCT 3120
Db 3061 AGATTAGCTCGCTCTGAGGAGGCGCTTACAGCGGCTTCTTCTGCTGAGCTTTTCTCT 3120
QY 3121 AGATGCTGTCTGCCATTTCTTCCAAAGTGACTTCTTATAAATCAAACTCTCTCTGCGACAG 3180
Db 3121 AGATGCTGTCTGCCATTTCTTCCAAAGTGACTTCTTATAAATCAAACTCTCTCTGCGACAG 3180
QY 3181 GCGGGAGAGCCAAATATAGACTTGTGTGAGCGCCCTACCTTGGGGGCTTTCCACG 3240
Db 3181 GCGGGAGAGCCAAATATAGACTTGTGTGAGCGCCCTACCTTGGGGGCTTTCCACG 3240
QY 3241 AGCTTGGGGAGAGCCATGATCTGAAATATAGTATATTTTGTAAATACGTGAACAA 3300
Db 3241 AGCTTGGGGAGAGCCATGATCTGAAATATAGTATATTTTGTAAATACGTGAACAA 3300
QY 3301 ACCAAACCCGTTTTTGTCTAAGGAGAAAGTAAATATGATTTTAAAAATCTATGTTTAA 3360
Db 3301 ACCAAACCCGTTTTTGTCTAAGGAGAAAGTAAATATGATTTTAAAAATCTATGTTTAA 3360
QY 3361 AATACTATGTAATTTTCTATCTATTTAGTATATTTTATGATGAAATAAACTTTC 3420
Db 3361 AATACTATGTAATTTTCTATCTATTTAGTATATTTTATGATGAAATAAACTTTC 3420
QY 3421 TACTGTAAAAAATAAAAAAAAAAAAAAAAAAAAAA 3453
Db 3421 TACTGTAAAAAATAAAAAAAAAAAAAAAAAAAAAA 3453

```

```

RESULT 3
US-07-946-507-1
; Sequence 1, Application US/07946507
; Patent No. 5283354
; GENERAL INFORMATION:
; APPLICANT: Lemischka, Ihor R.
; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL
; TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCLONE SYSTEMS INCORPORATED
; STREET: 180 VARICK STREET
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: U.S.A.
; ZIP: 10014
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/946,507
; FILING DATE: 19920917

```

```

/ CLASSIFICATION: 536
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US/07/813,593
/ FILING DATE: 24-DEC-1991
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US/07/793,065
/ FILING DATE: 15-NOV-1991
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US/07/728,913
/ FILING DATE: 28-JUN-1991
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US/07/679,666
/ FILING DATE: 02-APR-1991
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Feit, Irving N.
/ REGISTRATION NUMBER: 28,601
/ REFERENCE/DOCKET NUMBER: LEM-3-PPP
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 212-645-1405
/ TELEFAX: 212-645-2054
/ INFORMATION FOR SEQ ID NO. 1:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 3453 base pairs
/ TYPE: NUCLEIC ACID
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: cDNA
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: 31..3009
/ FEATURE:
/ NAME/KEY: mat peptide
/ LOCATION: 31..3006
/ US-07-946-507-1

```

```

Query Match      100.0%; Score 3453; DB 1; Length 3453;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3453; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCGCCCTGGCTACCGCGCGCTCCGGAGGCCATGCGGGCGTTGGCGCAGCGAGCGGCGG 60
DB 1 CGCGCCCTGGCTACCGCGCGCTCCGGAGGCCATGCGGGCGTTGGCGCAGCGAGCGGCGG 60

QY 61 CGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120
DB 61 CGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120

QY 121 CTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
DB 121 CTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180

QY 181 CCATCATCGTACCGAATGGTGGCGAGGATCCCGAGAACCTTCCAGTGTACCCCGAGGCGC 240
DB 181 CCATCATCGTACCGAATGGTGGCGAGGATCCCGAGAACCTTCCAGTGTACCCCGAGGCGC 240

QY 241 CAGAGTGAAGGACGGTATATGAGCGGCCACCGTGGAGTGGCGCGAGTCTGGGTCCATC 300
DB 241 CAGAGTGAAGGACGGTATATGAGCGGCCACCGTGGAGTGGCGCGAGTCTGGGTCCATC 300

QY 301 ACCCTGCAAGTCAGTCTGCGCACCCAGGCGGACCTTTCTGCGCTCTGGGTCTTTAGCAC 360
DB 301 ACCCTGCAAGTCAGTCTGCGCACCCAGGCGGACCTTTCTGCGCTCTGGGTCTTTAGCAC 360

QY 361 AGCTCCCTGGGTGCGCAGCGCCTTTGATTTTACAAACAGAGGAATCGTTTCCATGGCC 420
DB 361 AGCTCCCTGGGTGCGCAGCGCCTTTGATTTTACAAACAGAGGAATCGTTTCCATGGCC 420

QY 421 ATCTTGAACGTGACAGACCCAGGACGAGAGATACCTTCCATATTCAGACGGAACGC 480
DB 421 ATCTTGAACGTGACAGACCCAGGACGAGAGATACCTTCCATATTCAGACGGAACGC 480

QY 481 GCCAACTACACAGTACTGTTTACAGTGAATGTAAAGATACACAGCTGTATGCTTAAGG 540
DB 481 GCCAACTACACAGTACTGTTTACAGTGAATGTAAAGATACACAGCTGTATGCTTAAGG 540

```

```

DB 481 GCCAACTACACAGTACTGTTTACAGTGAATGTAAAGATACACAGCTGTATGCTTAAGG 540
QY 541 AGACCTTACTTTAGGAAGATGAAAAACAGGATGCACTCTCTGCACTCTCCGAGGGTGT 600
DB 541 AGACCTTACTTTAGGAAGATGAAAAACAGGATGCACTCTCTGCACTCTCCGAGGGTGT 600
QY 601 CCGGAGCCACCTGTGGAGTGGGTGCTCTGCACTCCCAAGGAAAGCTGTAAAGAAAGAA 660
DB 601 CCGGAGCCACCTGTGGAGTGGGTGCTCTGCACTCCCAAGGAAAGCTGTAAAGAAAGAA 660
QY 661 GGCCTCTGCTGTTCTCAGAAAGGAGGAAAGTACTTTCATGAGTGTTCGGAACAGACATC 720
DB 661 GGCCTCTGCTGTTCTCAGAAAGGAGGAAAGTACTTTCATGAGTGTTCGGAACAGACATC 720
QY 721 AGATGCTGTGTAGAAATGCACTGGCGCGCAATGCACCAAGCTGTTCACCATAGATCTA 780
DB 721 AGATGCTGTGTAGAAATGCACTGGCGCGCAATGCACCAAGCTGTTCACCATAGATCTA 780
QY 781 AACCAAGCTCTCTCAGAGCACACTGCCCCAGTTATTCTTGAAGTGGGGAAACCTTTGTCG 840
DB 781 AACCAAGCTCTCTCAGAGCACACTGCCCCAGTTATTCTTGAAGTGGGGAAACCTTTGTCG 840
QY 841 ATCAGGTGTAAGGCCATCCATGTGAACCATGATTCGGGCTCCTCTGGAGTGTGAAGAC 900
DB 841 ATCAGGTGTAAGGCCATCCATGTGAACCATGATTCGGGCTCCTCTGGAGTGTGAAGAC 900
QY 901 AAAGCCCTGGAGAGGGCAGCTACTTTGAGATGAGTACTTCCACAAACAGGACCATG 960
DB 901 AAAGCCCTGGAGAGGGCAGCTACTTTGAGATGAGTACTTCCACAAACAGGACCATG 960
QY 961 ATTCCGATTTCTTTGGCTTTGTTGTTCTTCCGTGGGAAGGAACGACACCGGATATTACACC 1020
DB 961 ATTCCGATTTCTTTGGCTTTGTTGTTCTTCCGTGGGAAGGAACGACACCGGATATTACACC 1020
QY 1021 TGCTTTCTCTCAAGACACCCAGCCAGTACAGCTGTGGTGACCATCTTGAAGAAAGGGTTT 1080
DB 1021 TGCTTTCTCTCAAGACACCCAGCCAGTACAGCTGTGGTGACCATCTTGAAGAAAGGGTTT 1080
QY 1081 ATAAACGCTACCACTCGCAAGAGAGTATGAATTCACCCGCTAGCAAGAAAGTCTGCTTC 1140
DB 1081 ATAAACGCTACCACTCGCAAGAGAGTATGAATTCACCCGCTAGCAAGAAAGTCTGCTTC 1140
QY 1141 TCAGTCAAGTTTAAAGCGTATCCCAAGATCCGATGCACTGCTGATCTTCTCTCAAGCTCA 1200
DB 1141 TCAGTCAAGTTTAAAGCGTATCCCAAGATCCGATGCACTGCTGATCTTCTCTCAAGCTCA 1200
QY 1201 TTTCTTGTGAACAGAGAGCGCTGGAGATGGGTACAGCATATCTAAATTTTGGATCAT 1260
DB 1201 TTTCTTGTGAACAGAGAGCGCTGGAGATGGGTACAGCATATCTAAATTTTGGATCAT 1260
QY 1261 AAGAACAGCCAGGAGAGTACATATTTATGCAAGAAATGATGACGCCAGTTCACCAAA 1320
DB 1261 AAGAACAGCCAGGAGAGTACATATTTATGCAAGAAATGATGACGCCAGTTCACCAAA 1320
QY 1321 ATGTTCAAGCTGAATATAAGAAAGAAACCTTCAAGTGTAGCAAAATGATGACGCCAG 1380
DB 1321 ATGTTCAAGCTGAATATAAGAAAGAAACCTTCAAGTGTAGCAAAATGATGACGCCAG 1380
QY 1381 GCGTCTGTTCTCTGATGGCTACCCGCTACCTCTTGGACCTTGAAGAGTGTTCGGAC 1440
DB 1381 GCGTCTGTTCTCTGATGGCTACCCGCTACCTCTTGGACCTTGAAGAGTGTTCGGAC 1440
QY 1441 AAATCTCCCAATTTGCAAGGAGGAAATCCCAAGAGGAGTTTGAATATAAGAGGTAAACAGA 1500
DB 1441 AAATCTCCCAATTTGCAAGGAGGAAATCCCAAGAGGAGTTTGAATATAAGAGGTAAACAGA 1500
QY 1501 AAAGTGTTTGGCGCAGTGGGTGTCGACAGTACTCTAAATATGAGTGAAGCGGGAAGGG 1560
DB 1501 AAAGTGTTTGGCGCAGTGGGTGTCGACAGTACTCTAAATATGAGTGAAGCGGGAAGGG 1560
QY 1561 CTTCTGGTCAAAATGCTGTCGCTACAAATTTCTATGGGACCGTCTTGGCAAAACCATCTTTTA 1620
DB 1561 CTTCTGGTCAAAATGCTGTCGCTACAAATTTCTATGGGACCGTCTTGGCAAAACCATCTTTTA 1620

```



```
QY 1621 AACTACAGGCGCCCTTCCCTTTCATCCAGACAAATCTCTCTATGCGACCATTTGGG 1680
Db 1621 AACTACAGGCGCCCTTCCCTTTCATCCAGACAAATCTCTCTATGCGACCATTTGGG 1680
QY 1681 CTCTGCTCCCTTTCATTTGTTTCTCATTTGTTGTTGATCTGCCACAAATACAAAAAGCAA 1740
Db 1681 CTCTGCTCCCTTTCATTTGTTTCTCATTTGTTGTTGATCTGCCACAAATACAAAAAGCAA 1740
QY 1741 TTTAGTACGAGAGTCAGCTGCGAGATGATCCAGGTGATCGGCCCTCTGGATACGAGTAC 1800
Db 1741 TTTAGTACGAGAGTCAGCTGCGAGATGATCCAGGTGATCGGCCCTCTGGATACGAGTAC 1800
QY 1801 TTCTAGCTTACCTTACGGGACTATGATATGATCTTAAGTGGGATTTCCCGAGAGAGAAC 1860
Db 1801 TTCTAGCTTACCTTACGGGACTATGATATGATCTTAAGTGGGATTTCCCGAGAGAGAAC 1860
QY 1861 TTAGAGTTTGGGAAGGTCCTGGGCTCTGGGCTTTTCGGGAGGGTGTATGAACGCCACCGGCC 1920
Db 1861 TTAGAGTTTGGGAAGGTCCTGGGCTCTGGGCTTTTCGGGAGGGTGTATGAACGCCACCGGCC 1920
QY 1921 TATGGCATTAGTAAACCGGAGTCTCAATTCAGGTGGCGGTGAAGATGCTTAAAGAGAAA 1980
Db 1921 TATGGCATTAGTAAACCGGAGTCTCAATTCAGGTGGCGGTGAAGATGCTTAAAGAGAAA 1980
QY 1981 GCTGACAGCTGTGAAAAGAGCTCTCATGTCCGAGCTCAAAATGATGACCCACCTGGGA 2040
Db 1981 GCTGACAGCTGTGAAAAGAGCTCTCATGTCCGAGCTCAAAATGATGACCCACCTGGGA 2040
QY 2041 CACCATGACAACTCGTGAATCTGCTGGGCGATGCACCTGTCCAGGGCGAGTGTACTTGT 2100
Db 2041 CACCATGACAACTCGTGAATCTGCTGGGCGATGCACCTGTCCAGGGCGAGTGTACTTGT 2100
QY 2101 ATTTTGAATATTTGCTATGTGTGACCTCTCTCAACTACCTTAAGAGTAAAGAGAGAA 2160
Db 2101 ATTTTGAATATTTGCTATGTGTGACCTCTCTCAACTACCTTAAGAGTAAAGAGAGAA 2160
QY 2161 TTTACAGGACATGGACAGAGATTTTAAAGAAATATTTTCAAGTCTTACCCCTACTTTC 2220
Db 2161 TTTACAGGACATGGACAGAGATTTTAAAGAAATATTTTCAAGTCTTACCCCTACTTTC 2220
QY 2221 CAGGCATATCAATTCAGCATGCTGTTTACGAGAGTTCAGTTACACCCCGCTTGT 2280
Db 2221 CAGGCATATCAATTCAGCATGCTGTTTACGAGAGTTCAGTTACACCCCGCTTGT 2280
QY 2281 GATCAGCTCTCAGGTTCAATGGGAATTCATTTCTGAAGATGAGATTAATATGAA 2340
Db 2281 GATCAGCTCTCAGGTTCAATGGGAATTCATTTCTGAAGATGAGATTAATATGAA 2340
QY 2341 AACAGAGAGGCTGGCAGAGAGAGAGAGAGATTTGAACTGTGCTGACGTTTGAAGAC 2400
Db 2341 AACAGAGAGGCTGGCAGAGAGAGAGAGAGATTTGAACTGTGCTGACGTTTGAAGAC 2400
QY 2401 CTCCTTTGCTTGGTACCAAGTGGCCAAAGCATGGAATTCCTGGAGTCAAGTCGTTGT 2460
Db 2401 CTCCTTTGCTTGGTACCAAGTGGCCAAAGCATGGAATTCCTGGAGTCAAGTCGTTGT 2460
QY 2461 GTCCACAGAGACCTGGCAGCAGGAATGTGTGTTGTCACCCAGGAGGTGGTGAAGATC 2520
Db 2461 GTCCACAGAGACCTGGCAGCAGGAATGTGTGTTGTCACCCAGGAGGTGGTGAAGATC 2520
QY 2521 TGTGACTTTGGACTGGCCCGAGACATCTCTGAGCGACTCCAGCTACGTCGAGGGGCAAC 2580
Db 2521 TGTGACTTTGGACTGGCCCGAGACATCTCTGAGCGACTCCAGCTACGTCGAGGGGCAAC 2580
QY 2581 GCACGGCTGCGGTGAAGTGTGATGTCACCCGAGAGCTTATTTGAAGGGATCTACAAATC 2640
Db 2581 GCACGGCTGCGGTGAAGTGTGATGTCACCCGAGAGCTTATTTGAAGGGATCTACAAATC 2640
QY 2641 AAGAGTACGCTGCTGCTCCTACGGCATCTCTCTGGGAGATATTTTCACTGGGTGTGAAC 2700
Db 2641 AAGAGTACGCTGCTGCTCCTACGGCATCTCTCTGGGAGATATTTTCACTGGGTGTGAAC 2700
```

```
QY 2701 CCTTACCTGGCAATCCTGTGACGCTAACTTCTATAAACTGATTCAGAGTGAATTTAAA 2760
Db 2701 CCTTACCTGGCAATCCTGTGACGCTAACTTCTATAAACTGATTCAGAGTGAATTTAAA 2760
QY 2761 ATGGAGAGCATTCTATGCCACAGAGAGGATATCTTTGTAATGCAATCCTGCTGGCT 2820
Db 2761 ATGGAGAGCATTCTATGCCACAGAGAGGATATCTTTGTAATGCAATCCTGCTGGCT 2820
QY 2821 TTTGACTCAAGGAAGCGGCATCTTCCCAACTGACTTCAATTTTAGATGTCAGCTG 2880
Db 2821 TTTGACTCAAGGAAGCGGCATCTTCCCAACTGACTTCAATTTTAGATGTCAGCTG 2880
QY 2881 GCAGAGCGAGAGAGCATGTATCAGAACTCATCCATCTACCAAAACAGGCGGCCCT 2940
Db 2881 GCAGAGCGAGAGAGCATGTATCAGAACTCATCCATCTACCAAAACAGGCGGCCCT 2940
QY 2941 CAGCAGAGAGCGGGCTCAGAGCCAGTCCCAAGCCAGGCGGAGGATTCACAGAGAA 3000
Db 2941 CAGCAGAGAGCGGGCTCAGAGCCAGTCCCAAGCCAGGCGGAGGATTCACAGAGAA 3000
QY 3001 AGAAGTTAGCGAGAGCGCTTGGACCCGCCCTAGCAGGCTGTAGACGCGAGAGCA 3060
Db 3001 AGAAGTTAGCGAGAGCGCTTGGACCCGCCCTAGCAGGCTGTAGACGCGAGAGCA 3060
QY 3061 AGATTAGCTCGCTCTGAGGAGCGCCCTACAGCGGCTTCTCGCTGGAATTTTCTCT 3120
Db 3061 AGATTAGCTCGCTCTGAGGAGCGCCCTACAGCGGCTTCTCGCTGGAATTTTCTCT 3120
QY 3121 AGATGCTGTGCTGATTAATCCAAAGTGACTTCTATAAAATCAAACTCTCTGCAAG 3180
Db 3121 AGATGCTGTGCTGATTAATCCAAAGTGACTTCTATAAAATCAAACTCTCTGCAAG 3180
QY 3181 GCGGAGAGCAATATGAGACTTGTGTTGAGCGCCCTACCTGCGGCGCTTTCACAG 3240
Db 3181 GCGGAGAGCAATATGAGACTTGTGTTGAGCGCGCTACCTGCGGCGCTTTCACAG 3240
QY 3241 AGCTTGAGGAGCAATGAGCTGATCTGAAATATAGTATATCTTGTAAATACGTGAACAA 3300
Db 3241 AGCTTGAGGAGCAATGAGCTGATCTGAAATATAGTATATCTTGTAAATACGTGAACAA 3300
QY 3301 ACCAAACCCGTTTTTCTAGGAGCAATATGATTTTAAATCTATGTTTAA 3360
Db 3301 ACCAAACCCGTTTTTCTAGGAGCAATATGATTTTAAATCTATGTTTAA 3360
QY 3361 AATGATGATTAATCTTCTATCTATTTAGTATATTTATGATGAAATTAATCTTC 3420
Db 3361 AATGATGATTAATCTTCTATCTATTTAGTATATTTATGATGAAATTAATCTTC 3420
QY 3421 TACTGTAAAAA 3453
Db 3421 TACTGTAAAAA 3453
```

RESULT 4

```
US-08-252-517-1
; Sequence 1, Application US/08252517
; Patent No. 5548065
; GENERAL INFORMATION:
; APPLICANT: Lemischka, Ihor R.
; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL
; TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESS: ImClone Systems Incorporated
; STREET: 180 Varick Street
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10014
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
```

SOFTWARE: PatentIn Release #1.0, Version #1.25

```
;;
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/252,517
;; FILING DATE: 31-OCT-1994
;; CLASSIFICATION: 530
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/977,451
;; FILING DATE: 19-NOV-1992
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/906,397
;; FILING DATE: 26-JUN-1992
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US PCT/US92/05401
;; FILING DATE: 28-JUN-1992
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: TW 81102961
;; FILING DATE: 15-APR-1992
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US PCT/US92/02750
;; FILING DATE: 02-APR-1992
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/813,593
;; FILING DATE: 24-DEC-1991
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/793,065
;; FILING DATE: 15-NOV-1991
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/728,913
;; FILING DATE: 28-JUN-1991
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/679,666
;; FILING DATE: 02-APR-1991
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Feit, Irving N.
;; REGISTRATION NUMBER: 28,601
;; REFERENCE/DOCKET NUMBER: LEM-3-7P
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 212-645-1405
;; TELEFAX: 212-645-2054
;; INFORMATION FOR SEQ ID NO: 1:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 3453 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: double
;; TOPOLOGY: linear
;; MOLECULE TYPE: cDNA
;; HYPOTHETICAL: NO
;; ANTI-SENSE: NO
;; FRAGMENT TYPE: N-terminal
;; FEATURE:
;; NAME/KEY: mat_peptide
;; LOCATION: 112..3006
;; FEATURE:
;; NAME/KEY: sig_peptide
;; LOCATION: 31..111
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: 31..3009
US-08-252-517-1

Query Match 100.0%; Score 3453; DB 1; Length 3453;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3453; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCGGCTGGCTACCGCGCTCGGAGGCGCATGCGGGCGTTGGCGAGCGCAGCGCGG 60
Db 1 GCGGCTGGCTACCGCGCTCGGAGGCGCATGCGGGCGTTGGCGAGCGCAGCGCGG 60
QY 61 CGGCTGCTGCTCTGTTGTTGTCAGTAATGATTCTTGAGCCGTTTACAAACCAAGAC 120
Db 61 CGGCTGCTGCTCTGTTGTTGTCAGTAATGATTCTTGAGCCGTTTACAAACCAAGAC 120
QY 121 CTGCTGTGATCAAGTGTGTTTAAATCAGTCATGAGAACAAATGGCTCATCAGCGGAAAG 180
```

Db 1201 TTTCTTGTGAACAGAGAGGCTGGAGGATGGGTACAGCATATCTAAATTTTGGCATCAT 1260
Qy 1261 AAGAACAGCCAGGAGGTACATATCTATGACAGAAATGATGACGCCAGTTCCACCAA 1320
Db 1261 AAGAACAGCCAGGAGGTACATATCTATGACAGAAATGATGACGCCAGTTCCACCAA 1320
Qy 1321 ATGTTACGCTGAATATAAGAAAGAAACCTCAAGTGTAGCAAAATGCTCAGCCAGCCAG 1380
Db 1321 ATGTTACGCTGAATATAAGAAAGAAACCTCAAGTGTAGCAAAATGCTCAGCCAGCCAG 1380
Qy 1381 GCGTCTGTCTCTGTATGAGTACCGGCTACCGCTACCGTCTTGGACCTGGAAGAGTTTCGGAC 1440
Db 1381 GCGTCTGTCTCTGTATGAGTACCGGCTACCGTCTTGGACCTGGAAGAGTTTCGGAC 1440
Qy 1441 AAATCTCCCAATGACCGAGAAATCCAGAGGAGTTTGGATATAAAGGCTTACAGA 1500
Db 1441 AAATCTCCCAATGACCGAGAAATCCAGAGGAGTTTGGATATAAAGGCTTACAGA 1500
Qy 1501 AAAGTGTGTCAGTGGGTGTGAGCAGTACTCTAAATATGAGTGGGCGCGGAAAGGG 1560
Db 1501 AAAGTGTGTCAGTGGGTGTGAGCAGTACTCTAAATATGAGTGGGCGCGGAAAGGG 1560
Qy 1561 CTTCCTGTCCTGTCGTCGTCACATCTATGCGCACGCTTTCGCAACCATCTTTTAA 1620
Db 1561 CTTCCTGTCCTGTCGTCGTCACATCTATGCGCACGCTTTCGCAACCATCTTTTAA 1620
Qy 1621 AACTCACAGGCCCCCTTCCCTTTCATGCAAGCAACATCTCTTATGCGACCATTTGG 1680
Db 1621 AACTCACAGGCCCCCTTCCCTTTCATGCAAGCAACATCTCTTATGCGACCATTTGG 1680
Qy 1681 CTCTCTCCCCCTTCAATGTTGTTCTCATTTGTTGATCTGCCACAAATACAAAAGCAA 1740
Db 1681 CTCTCTCCCCCTTCAATGTTGTTCTCATTTGTTGATCTGCCACAAATACAAAAGCAA 1740
Qy 1741 TTTAGGTACGAGTCTGAGATGATCCAGGTGACTTGGCCCCCTGGATTAACGAGTAC 1800
Db 1741 TTTAGGTACGAGTCTGAGATGATCCAGGTGACTTGGCCCCCTGGATTAACGAGTAC 1800
Qy 1801 TTTCTAGTTGACTTCAGGAGTATGAATATGACCTTAAAGTGGAGTTCCCGAGAGAA 1860
Db 1801 TTTCTAGTTGACTTCAGGAGTATGAATATGACCTTAAAGTGGAGTTCCCGAGAGAA 1860
Qy 1861 TTAGAGTTGGAGAGGTCTGGGTCTGGGCTTTTCGGGAGGTGATGACGCCAGGCC 1920
Db 1861 TTAGAGTTGGAGAGGTCTGGGTCTGGGCTTTTCGGGAGGTGATGACGCCAGGCC 1920
Qy 1921 TATGCAATTAGTAAACGGAGTCTCAATTCAGGTGGCGTGAAGATGCTTAAAGAGAAA 1980
Db 1921 TATGCAATTAGTAAACGGAGTCTCAATTCAGGTGGCGTGAAGATGCTTAAAGAGAAA 1980
Qy 1981 GCTGACGCTGTGAAGAGAGTCTCATGTGCGAGCTCAAAATGATGACCCACCTGGGA 2040
Db 1981 GCTGACGCTGTGAAGAGAGTCTCATGTGCGAGCTCAAAATGATGACCCACCTGGGA 2040
Qy 2041 CACCATGACAACTCGTGAATCTGCTGGGGGATGACACTGTACGGGCGAGTGTACTTG 2100
Db 2041 CACCATGACAACTCGTGAATCTGCTGGGGGATGACACTGTACGGGCGAGTGTACTTG 2100
Qy 2101 ATTTTGTGAATATGTTGCTATGTTGACCTCTCAACTACTAGAAAGTAAAGAGAGAAG 2160
Db 2101 ATTTTGTGAATATGTTGCTATGTTGACCTCTCAACTACTAGAAAGTAAAGAGAGAAG 2160
Qy 2161 TTTTCAGGACATGACAGAGATTTTAAAGGAACATAATTTTCAGTTCTTACCTACTTTTC 2220
Db 2161 TTTTCAGGACATGACAGAGATTTTAAAGGAACATAATTTTCAGTTCTTACCTACTTTTC 2220
Qy 2221 CAGGCACATTTCAATTCAGCATGCTGTTTACGAGAGTTTACCTTACCCCGCCCTTG 2280
Db 2221 CAGGCACATTTCAATTCAGCATGCTGTTTACGAGAGTTTACCTTACCCCGCCCTTG 2280
Qy 2281 GATCAGCTCTCAGGTTCAATGGAAATTCATTTCTTGAAGTGAATGATGAA 2340
Db 2281 GATCAGCTCTCAGGTTCAATGGAAATTCATTTCTTGAAGTGAATGATGAA 2340

Qy 2341 AACCAAGAGAGGCTGCAGAAAGAGAGGAGAAATTTGAACTGTCTGACGTTTGAAGAC 2400
Db 2341 AACCAAGAGAGGCTGCAGAAAGAGAGGAGAAATTTGAACTGTCTGACGTTTGAAGAC 2400
Qy 2401 CTCTCTTTGCTTTGCTGTACCAAGTGGCCAAAGGCATGGAATTCCTGAGTTCAAGTCTGT 2460
Db 2401 CTCTCTTTGCTTTGCTGTACCAAGTGGCCAAAGGCATGGAATTCCTGAGTTCAAGTCTGT 2460
Qy 2461 GTCCACAGAGACTGGCAGCAGGAGATGTGTTGTCACCCACGGGAAGTGTGGAAGATC 2520
Db 2461 GTCCACAGAGACTGGCAGCAGGAGATGTGTTGTCACCCACGGGAAGTGTGGAAGATC 2520
Qy 2521 TGTGACTTTTGGACTGGCCGAGACATCTCTGAGCGACTCTCAGCTACGTCGTCAGGGCAAC 2580
Db 2521 TGTGACTTTTGGACTGGCCGAGACATCTCTGAGCGACTCTCAGCTACGTCGTCAGGGCAAC 2580
Qy 2581 GCACGGCTGCGGTTGAAGTGGATGCGCACCCGAGAGCTTATTTGAAGGGATCTACACAATC 2640
Db 2581 GCACGGCTGCGGTTGAAGTGGATGCGCACCCGAGAGCTTATTTGAAGGGATCTACACAATC 2640
Qy 2641 AAGAGTGAAGTCTGCTTACCGCATCTTCTCTGGAGATATTTTCACTGGGTGTGAAC 2700
Db 2641 AAGAGTGAAGTCTGCTTACCGCATCTTCTCTGGAGATATTTTCACTGGGTGTGAAC 2700
Qy 2701 CCTTACCTGCGCATCTCTGTCGACGCTAACTTCTATAAACTGATTGAGTGTGATTTAAA 2760
Db 2701 CCTTACCTGCGCATCTCTGTCGACGCTAACTTCTATAAACTGATTGAGTGTGATTTAAA 2760
Qy 2761 ATGAGAGCAGCAATCTATGCCACAGAAAGGATATCTTTGTAATGAATCCTCTGCGGCT 2820
Db 2761 ATGAGAGCAGCAATCTATGCCACAGAAAGGATATCTTTGTAATGAATCCTCTGCGGCT 2820
Qy 2821 TTTGACTCAAGAAAGCGGCTTCTTCCCACTGACTTCAATTTTAGGATGTGAGCTG 2880
Db 2821 TTTGACTCAAGAAAGCGGCTTCTTCCCACTGACTTCAATTTTAGGATGTGAGCTG 2880
Qy 2881 GCAGAGGAGAGAGATGATCAGAACTCCATCCATCTACCAAAACAGGCGGCCCT 2940
Db 2881 GCAGAGGAGAGAGATGATCAGAACTCCATCCATCTACCAAAACAGGCGGCCCT 2940
Qy 2941 CAGCAGAGAGGCGGCTCAGAGCCAGTCCACAGCGCCAGGTGAAGATTTCAACAGAAA 3000
Db 2941 CAGCAGAGAGGCGGCTCAGAGCCAGTCCACAGCGCCAGGTGAAGATTTCAACAGAAA 3000
Qy 3001 AGAAGTTAGCAGAGGCGCTTGGACCCCGCCACTAGCAGGCTGTAGACCGCAGAGCCA 3060
Db 3001 AGAAGTTAGCAGAGGCGCTTGGACCCCGCCACTAGCAGGCTGTAGACCGCAGAGCCA 3060
Qy 3061 AGATTAGCTCGCTCTGAGGAAGCGCCCTACAGCGGCTGCTTGGTGGACTTTTCTCT 3120
Db 3061 AGATTAGCTCGCTCTGAGGAAGCGCCCTACAGCGGCTGCTTGGTGGACTTTTCTCT 3120
Qy 3121 AGATTAGCTCGCTCTGAGGAAGCGCCCTACAGCGGCTGCTTGGTGGACTTTTCTCT 3180
Db 3121 AGATTAGCTCGCTCTGAGGAAGCGCCCTACAGCGGCTGCTTGGTGGACTTTTCTCT 3180
Qy 3181 GCGGAGAGCCAAATATGAGACTTGTGTTGGTGGAGCCCGCTTACCCCTGGGGGCTTTTCAAG 3240
Db 3181 GCGGAGAGCCAAATATGAGACTTGTGTTGGTGGAGCCCGCTTACCCCTGGGGGCTTTTCAAG 3240
Qy 3241 AGCTTGAAGGGAAGCCATGTATCTGAAATATAGTATATCTTTGTAATACGTGAACAA 3300
Db 3241 AGCTTGAAGGGAAGCCATGTATCTGAAATATAGTATATCTTTGTAATACGTGAACAA 3300
Qy 3301 ACCAAACCCGTTTGTGCTAAGGGAAGCTTAATATGATTTTAAAACTATGTTTTAA 3360
Db 3301 ACCAAACCCGTTTGTGCTAAGGGAAGCTTAATATGATTTTAAAACTATGTTTTAA 3360
Qy 3361 AATACATGTAATCTTTTCTATCTATTTAGTATATTTTATGATGGAATAAATCTTTC 3420
Db 3361 AATACATGTAATCTTTTCTATCTATTTAGTATATTTTATGATGGAATAAATCTTTC 3420

1141 TCAGTCAGGTTTAAAGCGTATACCACGAATCCGATGCAGTGGATCTTCTCTCAAGCCTCA 1200
1141 TCAGTCAGGTTTAAAGCGTATACCACGAATCCGATGCAGTGGATCTTCTCTCAAGCCTCA 1200
1201 TTTCCTTTGTGAACAGAGAGGCGCTGGAGGATGGGTACAGCATATCTAAATTTTGGGATCAT 1260
1201 TTTCCTTTGTGAACAGAGAGGCGCTGGAGGATGGGTACAGCATATCTAAATTTTGGGATCAT 1260
1261 AAGAACAGCCAGGAGATCATATTTCTATGCAGAAATGATGACGCCAGTTTCAACAAA 1320
1261 AAGAACAGCCAGGAGATCATATTTCTATGCAGAAATGATGACGCCAGTTTCAACAAA 1320
1321 ATGTTCCAGCTGAATATAAGAAAGAAACCTCAAGTGTCTAGCAATGCTCAGCCAGCCAG 1380
1321 ATGTTCCAGCTGNATATAGAAAGAAACCTCAAGTGTCTAGCAATGCTCAGCCAGCCAG 1380
1381 GCGTCTGTTCTCTGATGGCTACCCGCTACCTCTTTGGACCTGGAAGAGTGTTCGGAC 1440
1381 GCGTCTGTTCTCTGATGGCTACCCGCTACCTCTTTGGACCTGGAAGAGTGTTCGGAC 1440
1441 AAATCTCCAAATTCACGGAGGAAATCCCAAGAGGAGTTTGGAAATAAAAAGGCTAACAGA 1500
1441 AAATCTCCAAATTCACGGAGGAAATCCCAAGAGGAGTTTGGAAATAAAAAGGCTAACAGA 1500
1501 AAAGTGTTTGGCCAGTGGGTGTCGACGAGTACTCTAAATATGATGAGGCGCGGAAGGG 1560
1501 AAAGTGTTTGGCCAGTGGGTGTCGAGGAGTACTCTAAATATGATGAGGCGCGGAAGGG 1560
1561 CTTCTGGTCAAAATGCTGTGCGTACAAATTCATATGGGCACGCTCTTGGCAAAACCATCTTTTAA 1620
1561 CTTCTGGTCAAAATGCTGTGCGTACAAATTCATATGGGCACGCTCTTGGCAAAACCATCTTTTAA 1620
1621 AACTACACAGGCGCCCTTCCCTTTTCATCCAAAGACAAATCTCTCTATGCGACCAATGGG 1680
1621 AACTACACAGGCGCCCTTCCCTTTTCATCCAAAGACAAATCTCTCTATGCGACCAATGGG 1680
1681 CTCGTCTCCCTTCAATTTGTTCTCATTTGTTGATCTGCGCACAAATACAAAAAGCAA 1740
1681 CTCGTCTCCCTTCAATTTGTTCTCATTTGTTGATCTGCGCACAAATACAAAAAGCAA 1740
1741 TTATAGTACGAGAGTCAGCTGCAGATGATCCAGTGCATCGGCCCTCGGATACAGAGTAC 1800
1741 TTATAGTACGAGAGTCAGCTGCAGATGATCCAGTGCATCGGCCCTCGGATACAGAGTAC 1800
1801 TTCTACGTTGACTTCAGGAGCTATGAATATGACCTTAAAGTGGAGTTCCCGAGAGAGAAC 1860
1801 TTCTACGTTGACTTCAGGAGCTATGAATATGACCTTAAAGTGGAGTTCCCGAGAGAGAAC 1860
1861 TTAGAGTTTGGAGAGTCTTGGGCTCGGGCTTTTCGGAGGGTGAATGAACGCCAGGCC 1920
1861 TTAGAGTTTGGAGAGTCTTGGGCTCGGGCTTTTCGGAGGGTGAATGAACGCCAGGCC 1920
1921 TATGGCATTAGTAAACGGAGTCTCAATTCAGTGGCGGTGAAGATGCTTAAAGAGAAA 1980
1921 TATGCAATTAGTAAACGGAGTCTCAATTCAGTGGCGGTGAAGATGCTTAAAGAGAAA 1980
1981 GCTGACAGCTGTGAAAGAGAGCTCTCATGTGGAGCTCAAAATGATGACCCACCTGGGA 2040
1981 GCTGACAGCTGTGAAAGAGAGCTCTCATGTGGAGCTCAAAATGATGACCCACCTGGGA 2040
2041 CACCATGACAACTGTTGAATCTGCTGGGGGATGACACTGTGAGGGCAGTGTACTTGG 2100
2041 CACCATGACAACTGTTGAATCTGCTGGGGGATGACACTGTGAGGGCAGTGTACTTGG 2100
2101 ATTTTGTGAATATTTGTTGATGCTGACCTCTCACTACTAAGAGATGAAGAGAGAGAG 2160
2101 ATTTTGTGAATATTTGTTGATGCTGACCTCTCACTACTAAGAGATGAAGAGAGAGAGAG 2160
2161 TTTCACAGGACATGGACAGAGATTTTAAAGGAACATAATTTTCAGTGTCTTACCCCTACTTTC 2220
2161 TTTCACAGGACATGGACAGAGATTTTAAAGGAACATAATTTTCAGTGTCTTACCCCTACTTTC 2220

2221 CAGGCACATTCAAATTCAGCATGSCCTGGTTTACGAGAAGTTTCAAGTTACACCGGCCCTTG 2280
2221 CAGGCACATTCAAATTCAGCATGSCCTGGTTTACGAGAAGTTTCAAGTTACACCGGCCCTTG 2280
2281 GATCAGCTCTCAGGTTTCAATGGGAATTCATTTCAATCTGGAAGATGAGATTGAATATGAA 2340
2281 GATCAGCTCTCAGGTTTCAATGGGAATTCATTTCAATCTGGAAGATGAGATTGAATATGAA 2340
2341 AACCAAGAGAGGCTGGCAGAAAGAGAGAGAGAGATTGAAACGTCGTGACGTTTGAAGAC 2400
2341 AACCAAGAGAGGCTGGCAGAAAGAGAGAGAGAGATTGAAACGTCGTGACGTTTGAAGAC 2400
2401 CTCCTTTGCTTTGCTGTACCAAGTGGCCAAAGGCATGGAATTCCTTGGAGTTCAAGTCTGT 2460
2401 CTCCTTTGCTTTGCTGTACCAAGTGGCCAAAGGCATGGAATTCCTTGGAGTTCAAGTCTGT 2460
2461 GTCACAGAGACCTGGCAGCCAGGAATGTGTGTGTCTACCCACGGGAAGGTGGTGAAGATC 2520
2461 GTCACAGAGACCTGGCAGCCAGGAATGTGTGTGTCTACCCACGGGAAGGTGGTGAAGATC 2520
2521 TGTGACTTTGGACTGGCCCGAGACATCTCTGAGCGACTCCAGTACGTCGTGAGGGGCAAC 2580
2521 TGTGACTTTGGACTGGCCCGAGACATCTCTGAGCGACTCCAGTACGTCGTGAGGGGCAAC 2580
2581 GCACGCTGCGGTGAAGTGGATGGCACCCGAGAGCTTATTTGAAGGATCTACACAAATC 2640
2581 GCACGCTGCGGTGAAGTGGATGGCACCCGAGAGCTTATTTGAAGGATCTACACAAATC 2640
2641 AAGAGTGAAGCTCTGGTCTACGGCATCTCTCTCTGGGAGATATTTTCACTGGGTGTGAAC 2700
2641 AAGAGTGAAGCTCTGGTCTACGGCATCTCTCTCTGGGAGATATTTTCACTGGGTGTGAAC 2700
2701 CCTTACCTGGCAATCTCTGTGACGCTAACTTCTATAAATGATTCAGAGTGGATTTAAA 2760
2701 CCTTACCTGGCAATCTCTGTGACGCTAACTTCTATAAATGATTCAGAGTGGATTTAAA 2760
2761 ATGAGCAGCCATCTATGCGCACAGAGGATATCTTGTGATGCAATCTCTGCTGGCT 2820
2761 ATGAGCAGCCATCTATGCGCACAGAGGATATCTTGTGATGCAATCTCTGCTGGCT 2820
2821 TTTGACTCAAGGAAGCGGCATCTTCCCAACCTGACTTCTATTTTAGGATGTCAGCTG 2880
2821 TTTGACTCAAGGAAGCGGCATCTTCCCAACCTGACTTCTATTTTAGGATGTCAGCTG 2880
2881 GCAGAGCGAAGAACATGATATCAGAACATCCATCTACCAAAACAGGCGGCCCT 2940
2881 GCAGAGCGAAGAACATGATATCAGAACATCCATCTACCAAAACAGGCGGCCCT 2940
2941 CAGCAGAGAGCGGCTCAGAGCCAGTGCACAGGCCAGGTGAAGATTACAGAGAA 3000
2941 CAGCAGAGAGCGGCTCAGAGCCAGTGCACAGGCCAGGTGAAGATTACAGAGAA 3000
3001 AGAAGTTACGAGGAGCGCTTGGACCCCGCCAGTGCAGAGGCTGTAGACCGCAGAGCCA 3060
3001 AGAAGTTACGAGGAGCGCTTGGACCCCGCCAGTGCAGAGGCTGTAGACCGCAGAGCCA 3060
3061 AGATTAGCCTCGCTCTGAGGAAGCGCCCTACAGCGCTTCTGCTGACATTTTCTCT 3120
3061 AGATTAGCCTCGCTCTGAGGAAGCGCCCTACAGCGGTTGCTTCTGCTGACATTTTCTCT 3120
3121 AGATGCTGTCTGCATTAATCTCAAAGTGAATCTTATAAATCAAACTCTCTCTCGCAGAG 3180
3121 AGATGCTGTCTGCATTAATCTCAAAGTGAATCTTATAAATCAAACTCTCTCTCGCAGAG 3180
3181 GCGGAGAGCCCAATATGAGACTTGTGTGAGCGCCCTACCCCTGGGGCTTTCACG 3240
3181 GCGGAGAGCCCAATATGAGACTTGTGTGAGCGCCCTACCCCTGGGGCTTTCACG 3240
3241 AGCTTGAAGGGAAGCCATGATCTGAATATAGTATATTTCTGTAAATACGTGAACAA 3300
3241 AGCTTGAAGGGAAGCCATGATCTGAATATAGTATATTTCTGTAAATACGTGAACAA 3300
3301 ACCAAACCGGTTTTTGTGTAAGGGAAGCTAAATATGATTTTTTAAAAATCTATGTTTAA 3360

Db	3301	ACCAACCCGTTTTTCTAGGGAAGCTAAATATGATTTTTAAAAATCTATGTTTAA	3360
Qy	3361	AATACTATGTAACCTTTTTTCATCTATTTAGTGATATATTTTTATGGATGGAATAAACCCTTC	3420
Db	3361	AATACTATGTAACCTTTTTTCATCTATTTAGTGATATATTTTTATGGATGGAATAAACCCTTC	3420
Qy	3421	TACTGTAAAAAaaaaaaaaaaaaaaaaaaaaa	3453
Db	3421	TACTGTAAAAAaaaaaaaaaaaaaaaaaaaaa	3453

RESULT 6
US-08-601-891-1
Sequence 1, Application US/08601891
Patent No. 5747651
GENERAL INFORMATION:
APPLICANT: Lemischka, Ihor R.
TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL
TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: ImClone Systems Incorporated
STREET: 180 Varick Street
CITY: New York
STATE: New York
COUNTRY: U.S.A.
Zip: 10014
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/601.891
FILING DATE: 15-FEB-1996
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/977,451
FILING DATE: 19-NOV-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/906,397
FILING DATE: 26-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US PCT/US92/05401
FILING DATE: 26-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: TW 81102961
FILING DATE: 15-APR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US PCT/US92/02750
FILING DATE: 02-APR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/813,593
FILING DATE: 24-DEC-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/793,065
FILING DATE: 15-NOV-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/728,913
FILING DATE: 28-JUN-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/679,666
FILING DATE: 02-APR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Feit, Irving N.
REGISTRATION NUMBER: 28,601
REFERENCE/DOCKET NUMBER: LEM-3-7P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-645-1405
TELEFAX: 212-645-2034
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:

Db 721 AGATGCTGTGTAGAAATGCACTGGGCGCGAATGACCAAGCTGTTTACCATAGATCTA 780
Qy 781 AACAGGCTCTCAGAGCACACTGCCCCAGTATTCTCTGAAAGTGGGGGAACCCCTTGTGG 840
Db 781 AACAGGCTCTCAGAGCACACTGCCCCAGTATTCTCTGAAAGTGGGGGAACCCCTTGTGG 840
Qy 841 ATCAGGTGTAAGGCCATCCATGTGAACCATGTGATTCGGGTCTCAGTGGGAGCTGGAAGAC 900
Db 841 ATCAGGTGTAAGGCCATCCATGTGAACCATGTGATTCGGGTCTCAGTGGGAGCTGGAAGAC 900
Qy 901 AAAGCCCTGGAGGAGGACGCTACTTTGAGATGAGTACTCTCCACAAACAGGACCATG 960
Db 901 AAAGCCCTGGAGGAGGACGCTACTTTGAGATGAGTACTCTCCACAAACAGGACCATG 960
Qy 961 ATTCCGATTTCTTGGCTTTGTCTTCGTGGGAAAGAACGACCGGATATTACACC 1020
Db 961 ATTCCGATTTCTTGGCTTTGTCTTCGTGGGAAAGAACGACCGGATATTACACC 1020
Qy 1021 TGTCTTCTCTCAAGACACCCAGCCAGTCTAGCGTTGGTGAACCATCTCTAGAAAAGGGTTT 1080
Db 1021 TGTCTTCTCTCAAGACACCCAGCCAGTCTAGCGTTGGTGAACCATCTCTAGAAAAGGGTTT 1080
Qy 1081 ATAAACGCTACCGCTCGCAAGAAGAGTATGAAATTGACCCGTCAGAAAAGTTCTGCTTC 1140
Db 1081 ATAAACGCTACCGCTCGCAAGAAGAGTATGAAATTGACCCGTCAGAAAAGTTCTGCTTC 1140
Qy 1141 TCAGTCAGGTTTAAAGGTCACCAAGATCCGATGCACTGGTATCTCTCAAGCCCTCA 1200
Db 1141 TCAGTCAGGTTTAAAGGTCACCAAGATCCGATGCACTGGTATCTCTCAAGCCCTCA 1200
Qy 1201 TTTCTCTGTGAACAGAGAGCCCTGGAGGATGGGTACAGCATATCTAAATTTTGGGATCAT 1260
Db 1201 TTTCTCTGTGAACAGAGAGCCCTGGAGGATGGGTACAGCATATCTAAATTTTGGGATCAT 1260
Qy 1261 AAGAAACAAGCCAGAGAGTACATATTCTATGCAGAAAATGATGACGCCACAGTTCACCAAA 1320
Db 1261 AAGAAACAAGCCAGAGAGTACATATTCTATGCAGAAAATGATGACGCCACAGTTCACCAAA 1320
Qy 1321 ATGTTCAAGCTGAATATAAGAAAGAAACCTCAAGTGTAGCAAAATGCTCAGCCAGCCAG 1380
Db 1321 ATGTTCAAGCTGAATATAAGAAAGAAACCTCAAGTGTAGCAAAATGCTCAGCCAGCCAG 1380
Qy 1381 GCGTCTGTCTCTGATGCTACCCGCTACCCCTCTTGGACCTGGAGAGAGTGTTCGGAC 1440
Db 1381 GCGTCTGTCTCTGATGCTACCCGCTACCCCTCTTGGACCTGGAGAGAGTGTTCGGAC 1440
Qy 1441 AAATCTCCCAATTTGCACGAGGAAATCCACAGAGGAGTTTGGAAATAAAAAGGCTTACAGA 1500
Db 1441 AAATCTCCCAATTTGCACGAGGAAATCCACAGAGGAGTTTGGAAATAAAAAGGCTTACAGA 1500
Qy 1501 AAAGTGTGTTGGCCAGTGGGTGTGAGCAGTACTCTAAATATGAGTGGAGCGGGAAGGG 1560
Db 1501 AAAGTGTGTTGGCCAGTGGGTGTGAGCAGTACTCTAAATATGAGTGGAGCGGGAAGGG 1560
Qy 1561 CTTCTGGTCAAAATGCTGTGGGTCAAAATCTATGGGACAGTCTTGGAAACCATCTTTTAA 1620
Db 1561 CTTCTGGTCAAAATGCTGTGGGTCAAAATCTATGGGACAGTCTTGGAAACCATCTTTTAA 1620
Qy 1621 AACTCAACGAGCCCTTCCCTTTTCATCCAGACAAACATCTCTTATGGACCATTTGGG 1680
Db 1621 AACTCAACGAGCCCTTCCCTTTTCATCCAGACAAACATCTCTTATGGACCATTTGGG 1680
Qy 1681 CTCTGTCTCCCTTCATTTGTTTCTCAITTTGTTGATCTGCCACAAATACAAAAGCAA 1740
Db 1681 CTCTGTCTCCCTTCATTTGTTTCTCAITTTGTTGATCTGCCACAAATACAAAAGCAA 1740
Qy 1741 TTTAGTACAGAGTCTGAGTGCAGATGATCCAGTGTACTGGCCCTCTGGATTAACAGTAC 1800
Db 1741 TTTAGTACAGAGTCTGAGTGCAGATGATCCAGTGTACTGGCCCTCTGGATTAACAGTAC 1800
Qy 1801 TTTCTAGTTTCACTTCAGGAGTATGATATGACCTTAAGTGGGAGTTCCCGAGAGAGAC 1860
Db 1801 TTTCTAGTTTCACTTCAGGAGTATGATATGACCTTAAGTGGGAGTTCCCGAGAGAGAC 1860

Qy 1861 TTAGAGTTTGGGAGGTCCTGGGGTCTGGCGCTTTTCGGGAGGGTGATGAACCGCACGACC 1920
Db 1861 TTAGAGTTTGGGAGGTCCTGGGGTCTGGCGCTTTTCGGGAGGGTGATGAACCGCACGACC 1920
Qy 1921 TATGGCAATTAGTAAAAACGGGAGTCTCAATTCAAGTGGCGGTGAAGATGCTTAAAAAGAGAA 1980
Db 1921 TATGGCAATTAGTAAAAACGGGAGTCTCAATTCAAGTGGCGGTGAAGATGCTTAAAAAGAGAA 1980
Qy 1981 GCTGACAGCTGTGAAAAAAGAGCTCTCATGTGGAGCTCAAAATGATGACCACTGGGA 2040
Db 1981 GCTGACAGCTGTGAAAAAAGAGCTCTCATGTGGAGCTCAAAATGATGACCACTGGGA 2040
Qy 2041 CACATGACAACTCGTGAATCTCTCGGGGCAATGACACACTGTGAGGGCCAGTGTACTTTG 2100
Db 2041 CACATGACAACTCGTGAATCTCTCGGGGCAATGACACACTGTGAGGGCCAGTGTACTTTG 2100
Qy 2101 ATTTTGAATATTGTTGCTATGTGTGACTCTCTCAACTACTCTAAAGAGTAAAAAGAGAG 2160
Db 2101 ATTTTGAATATTGTTGCTATGTGTGACTCTCTCAACTACTCTAAAGAGTAAAAAGAGAG 2160
Qy 2161 TTTTACAGGACATGACAGAGATTTTAAAGAAATATTTTCACTTCTTACCTACTTTTC 2220
Db 2161 TTTTACAGGACATGACAGAGATTTTAAAGAAATATTTTCACTTCTTACCTACTTTTC 2220
Qy 2221 CAGGCACATTTCAAAATTCAGCATGCTGTGTTTACAGAGAGTTCAGTTACACCGCCCTTG 2280
Db 2221 CAGGCACATTTCAAAATTCAGCATGCTGTGTTTACAGAGAGTTCAGTTACACCGCCCTTG 2280
Qy 2281 GATCAGCTCTCAGGGTTCAATGGGAATCAATTCATTTCTGAAGATGAGATTGAATATGAA 2340
Db 2281 GATCAGCTCTCAGGGTTCAATGGGAATCAATTCATTTCTGAAGATGAGATTGAATATGAA 2340
Qy 2341 AACAGAGAGGCTGGCAG 2400
Db 2341 AACAGAGAGGCTGGCAG 2400
Qy 2401 CTCCTTTGCTTTGGGTACCAAGTGGCCCAAGGCATGGAATTTCTTGGAGTTCAGTCTGT 2460
Db 2401 CTCCTTTGCTTTGGGTACCAAGTGGCCCAAGGCATGGAATTTCTTGGAGTTCAGTCTGT 2460
Qy 2461 GTCCACAGAGACCTGGCAGCCAGGAATGTGTGTCTACCCACGGAGAGTGTGTAAGATC 2520
Db 2461 GTCCACAGAGACCTGGCAGCCAGGAATGTGTGTCTACCCACGGAGAGTGTGTAAGATC 2520
Qy 2521 TGTGACTTTGGACTGGCCCGAGACATCTCTGAGCGACTCCAGTCTGTCAGGGGCAAC 2580
Db 2521 TGTGACTTTGGACTGGCCCGAGACATCTCTGAGCGACTCCAGTCTGTCAGGGGCAAC 2580
Qy 2581 GCACGGCTGCGGTGAAGTGGATGGACCCGAGAGCTTATTTGAAGGGATCTACCAATC 2640
Db 2581 GCACGGCTGCGGTGAAGTGGATGGACCCGAGAGCTTATTTGAAGGGATCTACCAATC 2640
Qy 2641 AAGAGTCACTCTGCTCTACCGCATCTTCTCTGGGAGATATTTTCACTGGGTGTGAAC 2700
Db 2641 AAGAGTCACTCTGCTCTACCGCATCTTCTCTGGGAGATATTTTCACTGGGTGTGAAC 2700
Qy 2701 CTTTACCTCTGGCATTTCTCTGTCAGCGCTAACTTTCTATAAACTGATTCAGAGTGGATTTAA 2760
Db 2701 CTTTACCTCTGGCATTTCTCTGTCAGCGCTAACTTTCTATAAACTGATTCAGAGTGGATTTAA 2760
Qy 2761 ATGAGCAGCCATTTATGCCCACAGAGGGATATCTTTGTAATGCAATCTCTGCTGGGCT 2820
Db 2761 ATGAGCAGCCATTTATGCCCACAGAGGGATATCTTTGTAATGCAATCTCTGCTGGGCT 2820
Qy 2821 TTTTGACTCAAGAGAGCGGCGCATCTTCCCAACCTGACTTCAATTTTAGGATGTGAGTGT 2880
Db 2821 TTTTGACTCAAGAGAGCGGCGCATCTTCCCAACCTGACTTCAATTTTAGGATGTGAGTGT 2880
Qy 2881 GCAGAGCAGAGAGAGATGTATCAGAACTTCCATCTATCCAAACACAGCGGCCCT 2940
Db 2881 GCAGAGCAGAGAGAGATGTATCAGAACTTCCATCTATCCAAACACAGCGGCCCT 2940


```

; FILING DATE: 15-APR-1992
; PRIORITY APPLICATION DATA: US PCT/US92/02750
; FILING DATE: 02-APR-1992
; PRIORITY APPLICATION DATA:
; FILING DATE: 24-DEC-1991
; FILING DATE: 07/813,593
; PRIORITY APPLICATION DATA:
; FILING DATE: 15-NOV-1991
; FILING DATE: 07/793,065
; PRIORITY APPLICATION DATA:
; FILING DATE: 28-JUN-1991
; FILING DATE: 07/728,913
; PRIORITY APPLICATION DATA:
; FILING DATE: 02-APR-1991
; FILING DATE: 07/679,666
; ATTORNEY/AGENT INFORMATION:
; NAME: Feit, Irving N.
; REGISTRATION NUMBER: 28,601
; REFERENCE/DOCKET NUMBER: LEM-3-7P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-645-1405
; TELEFAX: 212-645-2054
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3453 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 112..3006
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 31..111
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 31..3009
; US-09-021-324-1

Query Match 100.0%; Score 3453; DB 2; Length 3453;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3453; Conservative 0; Mismatches 0; Indels 0; Gaps 0

Qy 1 GCGGCCTGGCTACCGCGCGCTCCGGAGGCCATGCGGGCGTTGGCGCAGCGCAGCGACCGG 60
Db 1 GCGGCCTGGCTACCGCGCGCTCCGGAGGCCATGCGGGCGTTGGCGCAGCGCAGCGACCGG 60
Qy 61 CGGCTGCTGCTGCTGTTGTTGTTGTCAGTAATGATCTTGAGACCGTTTCAAAACCAAGAC 120
Db 61 CGGCTGCTGCTGCTGTTGTTGTTGTCAGTAATGATCTTGAGACCGTTTCAAAACCAAGAC 120
Qy 121 CTGCGCTGTGATCAAGTGTGTTTTTAATCAGTCATGAGAACCAATGCTCATCAGCGGGAAG 180
Db 121 CTGCGCTGTGATCAAGTGTGTTTTTAATCAGTCATGAGAACCAATGCTCATCAGCGGGAAG 180
Qy 181 CCATCATCGTACCGAATGGTGCAGGATCCCGAGAGACCTCCAGTGTACCCCGAGGCGC 240
Db 181 CCATCATCGTACCGAATGGTGCAGGATCCCGAGAGACCTCCAGTGTACCCCGAGGCGC 240
Qy 241 CAGAGTGAAGGGACGGTATATGAGCGGCCACCGTGTGAGTGGCCGAGTCTGGGTCCATC 300
Db 241 CAGAGTGAAGGGACGGTATATGAGCGGCCACCGTGTGAGTGGCCGAGTCTGGGTCCATC 300
Qy 301 ACCCTGCAAGTGCAGCTCGCCACCCAGGGGACCTTTCTCGTCCCTCTGGGTCTTTAGCAC 360
Db 301 ACCCTGCAAGTGCAGCTCGCCACCCAGGGGACCTTTCTCGTCCCTCTGGGTCTTTAGCAC 360
Qy 361 AGCTCCCTGGGTGCGCCAGCGGCATTTTGATTTTCAAAAACAGAGGAATCGTTTTCCATGGCC 420

```

RESOLVED

US-09/021-324-1

/ Sequence 1, Application US/09021324

/ Patent No. 5912133

/ GENERAL INFORMATION:

/ APPLICANT: Lenischka, Ihor R.

/ TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL

/ TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS

/ NUMBER OF SEQUENCES: 10

/ CORRESPONDENCE ADDRESS:

/ ADDRESSEE: Imclone Systems Incorporated

/ STREET: 180 Varick Street

/ CITY: New York

/ STATE: New York

/ COUNTRY: U.S.A.

/ ZIP: 10014

/ COMPUTER READABLE FORM:

/ MEDIUM TYPE: Floppy disk

/ COMPUTER: IBM PC compatible

/ OPERATING SYSTEM: PC-DOS/MS-DOS

/ SOFTWARE: PatentIn Release #1.0, Version #1.25

/ CURRENT APPLICATION DATA:

/ APPLICATION NUMBER: US/09/021.324

/ FILING DATE:

/ CLASSIFICATION:

/ PRIOR APPLICATION DATA:

/ APPLICATION NUMBER: US/07/977.451

/ FILING DATE: 1992-11-19

/ PRIOR APPLICATION DATA:

/ APPLICATION NUMBER: US 07/906.397

/ FILING DATE: 26-JUN-1992

/ PRIOR APPLICATION DATA:

/ APPLICATION NUMBER: US PCT/US92/05401

/ FILING DATE: 26-JUN-1992

/ PRIOR APPLICATION DATA:

/ APPLICATION NUMBER: TW 81102961

.

Db 361 ||||| AGCTCCCTGGCTGCCCCAGCCGCACTTTGATTTTACAAAACAGAGAAATCGTTTCCATGGCC 420
Qy 421 ATCTTGAACTGTACAGAGACCCAGGACAGGAGAAATACCTACTCCATATTCAGAGCGAACGC 480
Db 421 ATCTTGAACTGTACAGAGACCCAGGACAGGAGAAATACCTACTCCATATTCAGAGCGAACGC 480
Qy 481 GCCAACTACACAGTATCTGTTTACAGTGAATGTATAGAGATACACAGCTGTATGCTTAAGG 540
Db 481 GCCAACTACACAGTATCTGTTTACAGTGAATGTATAGAGATACACAGCTGTATGCTTAAGG 540
Qy 541 AGACCTTACTTTAGGAGATGGAACCCAGATGCACTGCTCTGCACTCTCCGAGGGTGT 600
Db 541 AGACCTTACTTTAGGAGATGGAACCCAGATGCACTGCTCTGCACTCTCCGAGGGTGT 600
Qy 601 CCGGAGCCCACTGTGAGTGGGTGCTCTGCACTGCCACAGGGAAGCTGTAAAGAGAA 660
Db 601 CCGGAGCCCACTGTGAGTGGGTGCTCTGCACTGCCACAGGGAAGCTGTAAAGAGAA 660
Qy 661 GGCCCTGCTGTGTACAGAGGAGAAAGGTACTTCAATGAGTGTTCGGAAACAGATC 720
Db 661 GGCCCTGCTGTGTACAGAGGAGAAAGGTACTTCAATGAGTGTTCGGAAACAGATC 720
Qy 721 AGATGCTGTGTAGAAATGCACTGGCGCGCAATGCAACCAAGCTGTTCACCATAGATCTA 780
Db 721 AGATGCTGTGTAGAAATGCACTGGCGCGCAATGCAACCAAGCTGTTCACCATAGATCTA 780
Qy 781 AACCAAGCTCTCAGAGACACTGCCAGTATTTCTGAAAGTGGGGGAACCCCTTGTGG 840
Db 781 AACCAAGCTCTCAGAGACACTGCCAGTATTTCTGAAAGTGGGGGAACCCCTTGTGG 840
Qy 841 ATCAGGTGTAAAGCCATCCATGTGAACCAATGATTCGGGCTCACCTGGAGCTGGAAGAC 900
Db 841 ATCAGGTGTAAAGCCATCCATGTGAACCAATGATTCGGGCTCACCTGGAGCTGGAAGAC 900
Qy 901 AAAGCCCTGAGAGGAGGAGTACTTTGAGATGAGTACTTCCACAAACAGGACCATG 960
Db 901 AAAGCCCTGAGAGGAGGAGTACTTTGAGATGAGTACTTCCACAAACAGGACCATG 960
Qy 961 ATTCGGATTCCTTGGCCCTTTGTGTCCTCGGGAAGGAGACGACACCGGATATTACAC 1020
Db 961 ATTCGGATTCCTTGGCCCTTTGTGTCCTCGGGAAGGAGACGACACCGGATATTACAC 1020
Qy 1021 TGCTCTTCTCAAGACACCCAGCCAGTCAGCGTTGGTGACCATCTAGAAAAGGGTTT 1080
Db 1021 TGCTCTTCTCAAGACACCCAGCCAGTCAGCGTTGGTGACCATCTAGAAAAGGGTTT 1080
Qy 1081 ATAAACGCTACAGCTCGCAAGAGAGTATGAAATTTGACCGTACGAAAGTTCTGCTTC 1140
Db 1081 ATAAACGCTACAGCTCGCAAGAGAGTATGAAATTTGACCGTACGAAAGTTCTGCTTC 1140
Qy 1141 TCAGTCAGGTTTAAAGCGTACCCAGAAATCCGATCGATGATCTTCTCAAGCCCTCA 1200
Db 1141 TCAGTCAGGTTTAAAGCGTACCCAGAAATCCGATCGATGATCTTCTCAAGCCCTCA 1200
Qy 1201 TTTCTCTGTAACAGAGAGGCTCGGAGGATGGGTACAGCATATCTAAATTTTTCGATCAT 1260
Db 1201 TTTCTCTGTAACAGAGAGGCTCGGAGGATGGGTACAGCATATCTAAATTTTTCGATCAT 1260
Qy 1261 AAGAAACAGCCAGGAGAGTACATATTTCTATGAGAAATGATGAGCCGAGTTTACCBAA 1320
Db 1261 AAGAAACAGCCAGGAGAGTACATATTTCTATGAGAAATGATGAGCCGAGTTTACCBAA 1320
Qy 1321 ATGTTACGCTCAATATAGAAAGAAACCTCAAGTGTAGCAAAATGCCTCAGCCAGCCAG 1380
Db 1321 ATGTTACGCTCAATATAGAAAGAAACCTCAAGTGTAGCAAAATGCCTCAGCCAGCCAG 1380
Qy 1381 GCGTCTGTTCTCTGATGGCTACCGCTACCTCTTGGACCTGGAAGAGTGTTCGGAC 1440
Db 1381 GCGTCTGTTCTCTGATGGCTACCGCTACCTCTTGGACCTGGAAGAGTGTTCGGAC 1440
Qy 1441 AAATCTCCCAATTCACCGAGGAAATCCAGAGGAGTTTGAATATAAAGGCTTAACAGA 1500

1441 AAATCTCCCAATTCACCGAGGAAATCCAGAGGAGTTTGAATATAAAGGCTTAACAGA 1500
Qy 1501 AAATGTTTGGCCAGTGGGTGTCGACAGTACTCTAAATATATGATGAGTGGCCGGAAGGG 1560
Db 1501 AAATGTTTGGCCAGTGGGTGTCGACAGTACTCTAAATATATGATGAGTGGCCGGAAGGG 1560
Qy 1561 CTTCTGCTCAAAATGCTGTGCTTACAAATTCATATGGGCACTGCTTTCGGAACCATCTTTT 1620
Db 1561 CTTCTGCTCAAAATGCTGTGCTTACAAATTCATATGGGCACTGCTTTCGGAACCATCTTTT 1620
Qy 1621 AACTCACAGGCCCTTCCCTTTTCATCCAAAGCAAACTCTCTCTATGCGACCATTTGG 1680
Db 1621 AACTCACAGGCCCTTCCCTTTTCATCCAAAGCAAACTCTCTCTATGCGACCATTTGG 1680
Qy 1681 CTCTGCTCCCTTCTCATTTGTTTCTCATTTGTTGATCTGCGCAAAATATAAAGCA 1740
Db 1681 CTCTGCTCCCTTCTCATTTGTTTCTCATTTGTTGATCTGCGCAAAATATAAAGCA 1740
Qy 1741 TTTAGGTACAGAGTCACTGAGTCACTGAGTCACTGCGCCCTCTGGATTAACAGATAC 1800
Db 1741 TTTAGGTACAGAGTCACTGAGTCACTGAGTCACTGCGCCCTCTGGATTAACAGATAC 1800
Qy 1801 TTTCTAGCTTTCAGGAGCTATGAATATGAACCTTAAGTGGGAGTTCCCGAGAGAGAAC 1860
Db 1801 TTTCTAGCTTTCAGGAGCTATGAATATGAACCTTAAGTGGGAGTTCCCGAGAGAGAAC 1860
Qy 1861 TTAGAGTTTGGGAAGGTCCTGGGCTTTCGGGCTTTCGGGAGGTTGATGAACCCACGCC 1920
Db 1861 TTAGAGTTTGGGAAGGTCCTGGGCTTTCGGGCTTTCGGGAGGTTGATGAACCCACGCC 1920
Qy 1921 TATGGCAATATGTAAGACGGGAGTCTCAATTCAGTGGCGGTGAAGATGCTAAAGAGAAA 1980
Db 1921 TATGGCAATATGTAAGACGGGAGTCTCAATTCAGTGGCGGTGAAGATGCTAAAGAGAAA 1980
Qy 1981 GCTGACAGCTGTGAAAAGAGAGCTCTCATGTGCGAGCTCAAAATGATGACCCACTGGGA 2040
Db 1981 GCTGACAGCTGTGAAAAGAGAGCTCTCATGTGCGAGCTCAAAATGATGACCCACTGGGA 2040
Qy 2041 CACCATGACAACTGCGTGAATCTGCTGGGGCATGCACTGTGAGGGCCAGTGTACTTTG 2100
Db 2041 CACCATGACAACTGCGTGAATCTGCTGGGGCATGCACTGTGAGGGCCAGTGTACTTTG 2100
Qy 2101 ATTTTGAATATGCTGCTATGCTGCTCACTCACTCACTCACTCACTCACTCACTCACTCA 2160
Db 2101 ATTTTGAATATGCTGCTATGCTGCTCACTCACTCACTCACTCACTCACTCACTCACTCA 2160
Qy 2161 TTTTCAAGGACATGCAAGAGATTTTAAAGGAAATATTTTCACTTCTTACCTACTTTC 2220
Db 2161 TTTTCAAGGACATGCAAGAGATTTTAAAGGAAATATTTTCACTTCTTACCTACTTTC 2220
Qy 2221 CAGGCACATTCAAATTCAGCATGCTGTTTCAAGAGGTTTCACTTACCCGCCCTTG 2280
Db 2221 CAGGCACATTCAAATTCAGCATGCTGTTTCAAGAGGTTTCACTTACCCGCCCTTG 2280
Qy 2281 GATCAGCTCTCAGGGTTCAATGGGAATTCATTCATCTGAGATGAGATTAATGAA 2340
Db 2281 GATCAGCTCTCAGGGTTCAATGGGAATTCATTCATCTGAGATGAGATTAATGAA 2340
Qy 2341 AACCAAGAGGCTGGCAAG 2400
Db 2341 AACCAAGAGGCTGGCAAG 2400
Qy 2401 CTCCTTCTTGGCTACCAAGTGGCCAAAGGATGGAATTCCTGGAGTTCAAGTCTGT 2460
Db 2401 CTCCTTCTTGGCTACCAAGTGGCCAAAGGATGGAATTCCTGGAGTTCAAGTCTGT 2460
Qy 2461 GTCCACAGAGACTGCGCAGCCAGGATGTTTGTTCACCCAGGAGAGGTTGAGATC 2520
Db 2461 GTCCACAGAGACTGCGCAGCCAGGATGTTTGTTCACCCAGGAGAGGTTGAGATC 2520
Qy 2521 TGTGATTTTGGACTGGCCCGGAGACATCTGAGGACTCCAGTCTGCTCAGGGGCAAC 2580
Db 2521 TGTGATTTTGGACTGGCCCGGAGACATCTGAGGACTCCAGTCTGCTCAGGGGCAAC 2580

```
QY 2581 GCACGGCTGCCGCTGAAGTGGATGCGACCCGAGAGCTTATTTGAAGGGATCTACACAATC 2640
Db 2581 GCACGGCTGCCGCTGAAGTGGATGCGACCCGAGAGCTTATTTGAAGGGATCTACACAATC 2640
QY 2641 AAGAGTGAGCTGTGCTCTACGGCATCTTCTCTGGAGATATTTTCACTGGGTGTGAAC 2700
Db 2641 AAGAGTGAGCTGTGCTCTACGGCATCTTCTCTGGAGATATTTTCACTGGGTGTGAAC 2700
QY 2701 CCTTACCTGGGCAATCTCTGTCAGCGTAATCTTCTATAAACTGATTCAGAGTGGATTTAA 2760
Db 2701 CCTTACCTGGGCAATCTCTGTCAGCGTAATCTTCTATAAACTGATTCAGAGTGGATTTAA 2760
QY 2761 ATGGAGCAGCCATCTATGCCACAGAGGATATCTTTGTAATCAATCTCTGGGT 2820
Db 2761 ATGGAGCAGCCATCTATGCCACAGAGGATATCTTTGTAATCAATCTCTGGGT 2820
QY 2821 TTTGACTCAAGGAAGCGGCATCTTCCCAACCTGACTTCTTTAGGATGTGAGCTG 2880
Db 2821 TTTGACTCAAGGAAGCGGCATCTTCCCAACCTGACTTCTTTAGGATGTGAGCTG 2880
QY 2881 GCAGAGCAGAGAGCATGTATCAGACATCCATCCATCTACCAAAACAGCGGCCCT 2940
Db 2881 GCAGAGCAGAGAGCATGTATCAGACATCCATCCATCTACCAAAACAGCGGCCCT 2940
QY 2941 CAGCAGAGAGCGGCTCAGAGCCAGTCGCCACAGCCAGGTGAAGTTCACAGAA 3000
Db 2941 CAGCAGAGAGCGGCTCAGAGCCAGTCGCCACAGCCAGGTGAAGTTCACAGAA 3000
QY 3001 AGAAGTTAGCAGGAGGCGCTTGGACCCCGCCACCTAGCAGGCTGTAGACCCAGAGCA 3060
Db 3001 AGAAGTTAGCAGGAGGCGCTTGGACCCCGCCACCTAGCAGGCTGTAGACCCAGAGCA 3060
QY 3061 AGATTAGCCTCGCTCTGAGGAGGCGCTCAGCGGCTGTGCTCGCTGCTTCT 3120
Db 3061 AGATTAGCCTCGCTCTGAGGAGGCGCTCAGCGGCTGTGCTCGCTGCTTCT 3120
QY 3121 AGATGCTGTGCTGCTTACTCCAAAGTCACTTATAAATCAACCTCTCTGCAAG 3180
Db 3121 AGATGCTGTGCTGCTTACTCCAAAGTCACTTATAAATCAACCTCTCTGCAAG 3180
QY 3181 GCGGAGAGCCAAATAGAGACTTGTGCTGAGCGCGCTTACCCTGCGGCGCTTCCACG 3240
Db 3181 GCGGAGAGCCAAATAGAGACTTGTGCTGAGCGCGCTTACCCTGCGGCGCTTCCACG 3240
QY 3241 AGCTTGGGGAAGAGCCATGATCTGAAATATGATATTTCTGTAATACGTGAACAA 3300
Db 3241 AGCTTGGGGAAGAGCCATGATCTGAAATATGATATTTCTGTAATACGTGAACAA 3300
QY 3301 ACCAAACCGTTTTTGTAGGGAAGCTAAATATGATTTTAAATCTATGTTTTAA 3360
Db 3301 ACCAAACCGTTTTTGTAGGGAAGCTAAATATGATTTTAAATCTATGTTTTAA 3360
QY 3361 AATACTATGTAACCTTTTCTATCTATTTAGTGATATTTTATGATGGAATAAATCTTC 3420
Db 3361 AATACTATGTAACCTTTTCTATCTATTTAGTGATATTTTATGATGGAATAAATCTTC 3420
QY 3421 TACTGTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 3453
Db 3421 TACTGTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 3453
```

RESULT 8

```
PCT-US92-02750-1
; Sequence 1, Application PC/IUS9202750
; GENERAL INFORMATION:
; APPLICANT: LEMISCHKA, IHOR R.
; TITLE OF INVENTION: Totipotent Hematopoietic Stem Cell
; NUMBER OF INVENTION: 8
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: IMCLONE SYSTEMS INCORPORATED
; STREET: 180 VARICK STREET
```

```
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: US
; ZIP: 10014
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/02750
; FILING DATE: 19920402
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: FEIT, IRVING N.
; REGISTRATION NUMBER: 28,601
; REFERENCE/DOCKET NUMBER: LEM-3-PPPT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-645-1405
; TELEFAX: 212-645-2054
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3453 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 31...3009
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 31...3006
; PCT-US92-02750-1
```

Query Match 100.0%; Score 3453; DB 5; Length 3453;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 3453; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 GCGGCTGTGCTGTACCGCGCTCCGAGGCCATGCGGGGCTTGGCGCAGCGAGCCACCGG 60
Db 1 GCGGCTGTGCTGTACCGCGCTCCGAGGCCATGCGGGGCTTGGCGCAGCGAGCCACCGG 60
QY 61 CGGCTGTGCTGTCTGTGTTTGTCTAGTAAATGATTTCTTGAGACCGTTACAAACCAAGAC 120
Db 61 CGGCTGTGCTGTCTGTGTTTGTCTAGTAAATGATTTCTTGAGACCGTTACAAACCAAGAC 120
QY 121 CTGCTGTGATCAAGTGTGTTTAAATCAGTCATGAGAAACAATGCTCATCAGCGGAAAG 180
Db 121 CTGCTGTGATCAAGTGTGTTTAAATCAGTCATGAGAAACAATGCTCATCAGCGGAAAG 180
QY 181 CCATCATCGTACCGAATGTCGAGGATCCCGAGAGCTCCCGAGAGCTCCAGTGTACCCGAGGCGC 240
Db 181 CCATCATCGTACCGAATGTCGAGGATCCCGAGAGCTCCCGAGAGCTCCAGTGTACCCGAGGCGC 240
QY 241 CAGAGTGAAGGAGCGGTATATGAAAGCGGCCACCGTGGAGGTGGCGGAGTCTGGGTCCATC 300
Db 241 CAGAGTGAAGGAGCGGTATATGAAAGCGGCCACCGTGGAGGTGGCGGAGTCTGGGTCCATC 300
QY 301 ACCCTGCAAGTGCAGCTCGCCACCCAGGGGAGCTTTCTGCTCTCTGGGTCTTTAAGCAC 360
Db 301 ACCCTGCAAGTGCAGCTCGCCACCCAGGGGAGCTTTCTGCTCTCTGGGTCTTTAAGCAC 360
QY 361 AGCTCCCTGGGCTGCCAGCGCATTTTGAATTTACAAAACAGAGAAATCGTTTCCATGCC 420
Db 361 AGCTCCCTGGGCTGCCAGCGCATTTTGAATTTACAAAACAGAGAAATCGTTTCCATGCC 420
QY 421 ATCTTGAACGTGACAGAGACCCAGGAGGAGATACCTACTTCCATATTTACAGAGCAAGC 480
Db 421 ATCTTGAACGTGACAGAGACCCAGGAGGAGATACCTACTTCCATATTTACAGAGCAAGC 480
QY 481 GCCAACTACACAGTCTGTTTCAAGTGAATGTAAAGATATACAGCTGTATGTCTAAGG 540
Db 481 GCCAACTACACAGTCTGTTTCAAGTGAATGTAAAGATATACAGCTGTATGTCTAAGG 540
```

481 GCCAATCACAGTACTGTTACAGTGAATGTAAGAGATACACAGCTGTATGTCTAAGG 540 Db
541 AGACCTTACTTTAGGAAGATGGAATAACAGATGCACTGTCTGTGATCTCCAGAGGTGTT 600 Qy
541 AGACCTTACTTTAGGAAGATGGAATAACAGATGCACTGTCTGTGATCTCCAGAGGTGTT 600 Db
601 CCGAGGCCACTGTGGAGTGGGTCTGTGAGCTCCACAGGGAAGAGCTGTAAGAAGAA 660 Qy
601 CCGAGGCCACTGTGGAGTGGGTCTGTGAGCTCCACAGGGAAGAGCTGTAAGAAGAA 660 Db
661 GGCCCTCTGTGTGAGGAAGAGGAAGAGTACTTTCATGAGTGTGTTCCGGAACAGATC 720 Qy
661 GGCCCTCTGTGTGAGGAAGAGGAAGAGTACTTTCATGAGTGTGTTCCGGAACAGATC 720 Db
721 AGATGCTGTGTAGAAATGCACTGGGCGCGAATGCAACCAAGCTGTTCACATGATCTA 780 Qy
721 AGATGCTGTGTAGAAATGCACTGGGCGCGAATGCAACCAAGCTGTTCACATGATCTA 780 Db
781 AACAGGCTCTCAGACACACTGCCCGAGTTATTCCTGAAAGTGGGGGAACTTGTGG 840 Qy
781 AACAGGCTCTCAGACACACTGCCCGAGTTATTCCTGAAAGTGGGGGAACTTGTGG 840 Db
841 ATCAGGTGTAAGGCCATCTCAATGCAACCATGCAATTCGGGCTCACTGGGAGCTGGAAGAC 900 Qy
841 ATCAGGTGTAAGGCCATCTCAATGCAACCATGCAATTCGGGCTCACTGGGAGCTGGAAGAC 900 Db
901 AAAGCCCTGGAGGAGGAGCTACTTTGAGTGTGATCTCACTCCACAAACAGGACCATG 960 Qy
901 AAAGCCCTGGAGGAGGAGCTACTTTGAGTGTGATCTCACTCCACAAACAGGACCATG 960 Db
961 ATTCCGATTTCTTTGGGCTTTGTGTCTTCCTGGGAGGAAGACGACCGGATATTACACC 1020 Qy
961 ATTCCGATTTCTTTGGGCTTTGTGTCTTCCTGGGAGGAAGACGACCGGATATTACACC 1020 Db
1021 TGCTCTTCTCAAGACACCCAGCCAGTCAGCGTTGGTGACCATCTCTAGAAAAGGGTTT 1080 Qy
1021 TGCTCTTCTCAAGACACCCAGCCAGTCAGCGTTGGTGACCATCTCTAGAAAAGGGTTT 1080 Db
1081 ATAAACGCTACACGCTCGCAAGAGAGATGAAATTTGACCGGTACGAAAGTTCTGCTTC 1140 Qy
1081 ATAAACGCTACACGCTCGCAAGAGAGATGAAATTTGACCGGTACGAAAGTTCTGCTTC 1140 Db
1141 TCAGTCAGGTTTAAAGGTCACCAAGATCCGATCGATGCGATGATCTCTCAAGCCTCA 1200 Qy
1141 TCAGTCAGGTTTAAAGGTCACCAAGATCCGATCGATGCGATGATCTCTCAAGCCTCA 1200 Db
1201 TTTTCTTGTGAACAGAGGCGCTGGAGGATGGGTACGATATCTAAATTTTGGATCAT 1260 Qy
1201 TTTTCTTGTGAACAGAGGCGCTGGAGGATGGGTACGATATCTAAATTTTGGATCAT 1260 Db
1261 AAGAAACAGCAGGAGAGTACATATTTCTATGCAAAAATGATGCGCCAGTTACACCAA 1320 Qy
1261 AAGAAACAGCAGGAGAGTACATATTTCTATGCAAAAATGATGAGCCAGTTACACCAA 1320 Db
1321 ATGTTCAAGCTGAATATAAGAAAGAAACCTCAAGTGTAGCAATGCTCAGCCAGCCAG 1380 Qy
1321 ATGTTCAAGCTGAATATAAGAAAGAAACCTCAAGTGTAGCAATGCTCAGCCAGCCAG 1380 Db
1381 GCGTCCTGTTCTCTGATGCGCTACCGCTACCTCTTGGACCTGGAAGAGTGTTCGGAC 1440 Qy
1381 GCGTCCTGTTCTCTGATGCGCTACCGCTACCTCTTGGACCTGGAAGAGTGTTCGGAC 1440 Db
1441 AAATCTCCCAATTGCAAGGAGAAATCCAGAGGAGTTTGGAAATATAAGGCTAACAGA 1500 Qy
1441 AAATCTCCCAATTGCAAGGAGAAATCCAGAGGAGTTTGGAAATATAAGGCTAACAGA 1500 Db
1501 AAAGTGTGTGCGAGTGGGTGTCAGAGCAGTACTCTAAATATGAGTGGAGCCGGGAAAGG 1560 Qy
1501 AAAGTGTGTGCGAGTGGGTGTCAGAGCAGTACTCTAAATATGAGTGGAGCCGGGAAAGG 1560 Db
1561 CTTCTGGTCAATGCTGTGCGTACAAATTTCTATGGCAGCTCTTGGGAAACCATCTTTT 1620 Qy
1561 CTTCTGGTCAATGCTGTGCGTACAAATTTCTATGGCAGCTCTTGGGAAACCATCTTTT 1620 Db

1621 AACTCACAGGCGCCCTTCCCTTTTCATCAACAGCAACATCTCTCTTATGCGACCACTTGG 1680 Qy
1621 AACTCACAGGCGCCCTTCCCTTTTCATCAACAGCAACATCTCTCTTATGCGACCACTTGG 1680 Db
1681 CTCTGTCTCCCTTTCAATTTGTTGTTCTCATTTGTTGATCTGCCCAAAATACAAAAGCAA 1740 Qy
1681 CTCTGTCTCCCTTTCAATTTGTTGTTCTCATTTGTTGATCTGCCCAAAATACAAAAGCAA 1740 Db
1741 TTTAGGTACGAGAGTACAGTGCAGATGATCAAGTGCAGTGGCCCTCGATACAGATAC 1800 Qy
1741 TTTAGGTACGAGAGTACAGTGCAGATGATCAAGTGCAGTGGCCCTCGATACAGATAC 1800 Db
1801 TTCTACGTTGACTTTCAGGACTATGAATATGACCTTAAAGTGGAGTTCCCGAGAGAAC 1860 Qy
1801 TTCTACGTTGACTTTCAGGACTATGAATATGACCTTAAAGTGGAGTTCCCGAGAGAAC 1860 Db
1861 TTAGAGTTTGGGAAGGCTCTGGGCTTGGGCTTTTCGGGAGGTTGATGAACCCACGGCC 1920 Qy
1861 TTAGAGTTTGGGAAGGCTCTGGGCTTGGGCTTTTCGGGAGGTTGATGAACCCACGGCC 1920 Db
1921 TATGGCAATTAGTAAACCGGAGTCTCAATTCAGTGGCGGTGAAGATGCTAAAAGAGAA 1980 Qy
1921 TATGGCAATTAGTAAACCGGAGTCTCAATTCAGTGGCGGTGAAGATGCTAAAAGAGAA 1980 Db
1981 GCTCAGAGTGTGAAAAGAGAGCTCTCATGTCGGAGCTCAAAATGATGACCCACCTGGA 2040 Qy
1981 GCTCAGAGTGTGAAAAGAGAGCTCTCATGTCGGAGCTCAAAATGATGACCCACCTGGA 2040 Db
2041 CACCATGACAACTCGTGAATCTGTCGGGCGATGCACACTGTCTAGGGCCAGTGTACTTG 2100 Qy
2041 CACCATGACAACTCGTGAATCTGTCGGGCGATGCACACTGTCTAGGGCCAGTGTACTTG 2100 Db
2101 ATTTTGTGAATTTGCTGTGTGACCTCTCAACTACCTAAGAAAGTAAAAGAGAGAG 2160 Qy
2101 ATTTTGTGAATTTGCTGTGTGACCTCTCAACTACCTAAGAAAGTAAAAGAGAGAG 2160 Db
2161 TTTTCAAGGACATGGAAGAGATTTTAAAGAAATATTTTCAAGTCTTACCTACTTTC 2220 Qy
2161 TTTTCAAGGACATGGAAGAGATTTTAAAGAAATATTTTCAAGTCTTACCTACTTTC 2220 Db
2221 CAGGCAATTTCAAAATTCAGCATGCTGTTTACAGAGAGTTCAGTTACACCCGCTTTC 2280 Qy
2221 CAGGCAATTTCAAAATTCAGCATGCTGTTTACAGAGAGTTCAGTTACACCCGCTTTC 2280 Db
2281 GATCAGCTCTCAGGCTTCAATGGGAATTTCAATTTCTGAAGATGAGATTTGAATGAA 2340 Qy
2281 GATCAGCTCTCAGGCTTCAATGGGAATTTCAATTTCTGAAGATGAGATTTGAATGAA 2340 Db
2341 AACCAAGAGGCTGGCGAAGAGAGAGAGAGATTTGAACGCTGCTGACGTTTGAAGAC 2400 Qy
2341 AACCAAGAGGCTGGCGAAGAGAGAGAGAGATTTGAACGCTGCTGACGTTTGAAGAC 2400 Db
2401 CTCCTTTGCTTGTGATACAGTGGCCAAAGGCAATGGAATTCCTGGAGTTCAGTCTGCT 2460 Qy
2401 CTCCTTTGCTTGTGATACAGTGGCCAAAGGCAATGGAATTCCTGGAGTTCAGTCTGCT 2460 Db
2461 GTCCACAGAGACCTGGCAGCCAGAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2520 Qy
2461 GTCCACAGAGACCTGGCAGCCAGAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2520 Db
2521 TGTGACTTTGACTGGCCGAGACATCTGAGCGACTCCAGCTACGCTGCTGAGGAGTAC 2580 Qy
2521 TGTGACTTTGACTGGCCGAGACATCTGAGCGACTCCAGCTACGCTGCTGAGGAGTAC 2580 Db
2581 GCACGGCTGCGGTTGAAGTGGATGGATGGATGGATGGATGGATGGATGGATGGATGG 2640 Qy
2581 GCACGGCTGCGGTTGAAGTGGATGGATGGATGGATGGATGGATGGATGGATGGATGG 2640 Db
2641 AAGAGTGAAGTGTGCTCTCAGCGCATCTCTCTGGGAGATATTTTCACTGGGTGTGAAC 2700 Qy
2641 AAGAGTGAAGTGTGCTCTCAGCGCATCTCTCTGGGAGATATTTTCACTGGGTGTGAAC 2700 Db

Db 661 GGCCCTGCTGTGTGTCAGAAAAGGAGAAAAGGTACTTTCATGAGTTGTTTCGGAACAGACATC 720
Qy 721 AGATGCTGTCTAGAAATGCACTGGCGCGGAATGACCAAGCTGTTCACCATAGATCTA 780
Db 721 AGATGCTGTCTAGAAATGCACTGGCGCGGAATGACCAAGCTGTTCACCATAGATCTA 780
Qy 781 AACCAGGCTCTCTCAGAGCACACTGCCCCAGTTATTCCTGAAAGTGGGGGAACCCCTTGTGG 840
Db 781 AACCAGGCTCTCTCAGAGCACACTGCCCCAGTTATTCCTGAAAGTGGGGGAACCCCTTGTGG 840
Qy 841 ATCAGGTGTAAGGCATCCATGTGAACCATGTGAATTCGGGCTCACCTGGGAGCTGGAAGAC 900
Db 841 ATCAGGTGTAAGGCATCCATGTGAACCATGTGAATTCGGGCTCACCTGGGAGCTGGAAGAC 900
Qy 901 AAAGCCCTGAGGAGGCGACCTACTTTGAGATGAGTACTTCTCCACAAACAGACACCATG 960
Db 901 AAAGCCCTGAGGAGGCGACCTACTTTGAGATGAGTACTTCTCCACAAACAGACACCATG 960
Qy 961 ATTCCGATTCCTTTGGGCTTTGTGCTTCCTGGGAGGAAGACACACCGGATATTACACC 1020
Db 961 ATTCCGATTCCTTTGGGCTTTGTGCTTCCTGGGAGGAAGACACACCGGATATTACACC 1020
Qy 1021 TGTCTTCTTCAAAGCACCCAGCCAGCTGACGCTGTGGTGAACCATCTCTAGAAAAGGGTTT 1080
Db 1021 TGTCTTCTTCAAAGCACCCAGCCAGCTGACGCTGTGGTGAACCATCTCTAGAAAAGGGTTT 1080
Qy 1081 ATAAAGCTACCAAGCTGCGAAGAGATATGAATTTGACCCGTACGAAAAGTTCTGCTTC 1140
Db 1081 ATAAAGCTACCAAGCTGCGAAGAGATATGAATTTGACCCGTACGAAAAGTTCTGCTTC 1140
Qy 1141 TCAGTCAGGTTTAAAGGTTAAAGGTTAAAGGTTAAAGGTTAAAGGTTAAAGGTTAAAGGTT 1200
Db 1141 TCAGTCAGGTTTAAAGGTTAAAGGTTAAAGGTTAAAGGTTAAAGGTTAAAGGTTAAAGGTT 1200
Qy 1201 TTTCTTGTGAACAGAGAGGCTGGAGGATGGGTACAGCATATCTAAATTTTGGATCAT 1260
Db 1201 TTTCTTGTGAACAGAGAGGCTGGAGGATGGGTACAGCATATCTAAATTTTGGATCAT 1260
Qy 1261 AAGAACAGCCAGGAGATGATATCTATGCAAGAAATGATGAGCCGCTGATCAACAAA 1320
Db 1261 AAGAACAGCCAGGAGATGATATCTATGCAAGAAATGATGAGCCGCTGATCAACAAA 1320
Qy 1321 ATGTTCAAGCTGATATGAAGAGAAACCTCAAGTGTAGCAATGCTCAGCCAGCCAG 1380
Db 1321 ATGTTCAAGCTGATATGAAGAGAAACCTCAAGTGTAGCAATGCTCAGCCAGCCAG 1380
Qy 1381 GCGTCTGTTCTCTGATGCTACCCGCTACCCCTCTTGGACCTTGGAGAAAGTGTTCGGAC 1440
Db 1381 GCGTCTGTTCTCTGATGCTACCCGCTACCCCTCTTGGACCTTGGAGAAAGTGTTCGGAC 1440
Qy 1441 AAATCTCCCAATTCACGAGGAAATCCCAAGAGGATTTGGAATAAAAAGGCTAACAGA 1500
Db 1441 AAATCTCCCAATTCACGAGGAAATCCCAAGAGGATTTGGAATAAAAAGGCTAACAGA 1500
Qy 1501 AAAGTGTTCGACAGTGGGTGTCAGACAGTACTCTAAATATGATGAGGCGGGAAGGG 1560
Db 1501 AAAGTGTTCGACAGTGGGTGTCAGACAGTACTCTAAATATGATGAGGCGGGAAGGG 1560
Qy 1561 CTTCTGTTCAAAATGCTGTGCTGCAATTTCTATGGGACAGCTTTCGGAACCATCTTTTA 1620
Db 1561 CTTCTGTTCAAAATGCTGTGCTGCAATTTCTATGGGACAGCTTTCGGAACCATCTTTTA 1620
Qy 1621 AACTCACCAGGCCCCCTTCCCTTTTCATCCAAAGACAACTCTCCTTCTATGGACCATTTGG 1680
Db 1621 AACTCACCAGGCCCCCTTCCCTTTTCATCCAAAGACAACTCTCCTTCTATGGACCATTTGG 1680
Qy 1681 CTCGTCTCCCTTCATGTTGTTCTCATTTGTTGATCTGCAACAAATACAAAAGCAA 1740
Db 1681 CTCGTCTCCCTTCATGTTGTTCTCATTTGTTGATCTGCAACAAATACAAAAGCAA 1740
Qy 1741 TTTAGGTACGAGAGTCAAGTGCAGATGATCCAGTGTACTGCCCCCTCGGATACGAGTAC 1800
Db 1741 TTTAGGTACGAGAGTCAAGTGCAGATGATCCAGTGTACTGCCCCCTCGGATACGAGTAC 1800

Qy 1801 TTCTACGTTGACTTTCAGGAGCTATGAAATGACCTTAAAGTGGAGTTCCCGAGAGAGAAC 1860
Db 1801 TTCTACGTTGACTTTCAGGAGCTATGAAATGACCTTAAAGTGGAGTTCCCGAGAGAGAAC 1860
Qy 1861 TTAGAGTTTGGGAAGGTCCTGGGGTCTGGCGCTTTTCGGGAGGGTGTAAACCCACGCC 1920
Db 1861 TTAGAGTTTGGGAAGGTCCTGGGGTCTGGCGCTTTTCGGGAGGGTGTAAACCCACGCC 1920
Qy 1921 TATGGCATTTAGTAAACCGGAGTCTCAATTCAGTGGCGGTGAAGATGCTAAAGAGAGAA 1980
Db 1921 TATGGCATTTAGTAAACCGGAGTCTCAATTCAGTGGCGGTGAAGATGCTAAAGAGAGAA 1980
Qy 1981 GCTGACAGCTGTCAAAAAGAGCTCTCATCTCGAGCTCAAAATGATGACCCACCTGGGA 2040
Db 1981 GCTGACAGCTGTCAAAAAGAGCTCTCATCTCGAGCTCAAAATGATGACCCACCTGGGA 2040
Qy 2041 CACCATGACAAATCTGTGAATCTGTGGGGCATGCACACTGTCCAGGCGCATGTACTTTG 2100
Db 2041 CACCATGACAAATCTGTGAATCTGTGGGGCATGCACACTGTCCAGGCGCATGTACTTTG 2100
Qy 2101 ATTTTGAATTTTGTCTATGTGACCTCTCAACTACCTAAAGAGTAAAGAGAGAG 2160
Db 2101 ATTTTGAATTTTGTCTATGTGACCTCTCAACTACCTAAAGAGTAAAGAGAGAG 2160
Qy 2161 TTTTCAAGGACATGGAAGATTTTAAAGAACATTAATTTTCACTTCTTACCTTCTTTC 2220
Db 2161 TTTTCAAGGACATGGAAGATTTTAAAGAACATTAATTTTCACTTCTTACCTTCTTTC 2220
Qy 2221 CAGGCACTTCAAAATTCAGCATGCTGTTTACGAGAGTTCAGTTTACACCCGCCCTTG 2280
Db 2221 CAGGCACTTCAAAATTCAGCATGCTGTTTACGAGAGTTCAGTTTACACCCGCCCTTG 2280
Qy 2281 GATCAGCTCTCAGGTTCAATGGGAATTCATTTCTGAAGATGAGATGATATGAA 2340
Db 2281 GATCAGCTCTCAGGTTCAATGGGAATTCATTTCTGAAGATGAGATGATATGAA 2340
Qy 2341 AACCAAGAGGCTGGCAGAAAGAGGAGGAAGATTTTGAACGTGCTGAGCTTTGAAGAC 2400
Db 2341 AACCAAGAGGCTGGCAGAAAGAGGAGGAAGATTTTGAACGTGCTGAGCTTTGAAGAC 2400
Qy 2401 CTCCTTGTCTTTCGTTACCAAGTGGCCAAAGGCATGGAATTCCTGGAGTTTCAAGTCTGT 2460
Db 2401 CTCCTTGTCTTTCGTTACCAAGTGGCCAAAGGCATGGAATTCCTGGAGTTTCAAGTCTGT 2460
Qy 2461 GTCCACAGAGACCTGGCAGCCAGGAATGTGTTGTTCAACCCACCGGAAGGTGGAAGATC 2520
Db 2461 GTCCACAGAGACCTGGCAGCCAGGAATGTGTTGTTCAACCCACCGGAAGGTGGAAGATC 2520
Qy 2521 TGTGACTTTTGGACTGCGCCGAGACATCTCTGAGCGACTCCAGCTACGTCGTCAGGGCAAC 2580
Db 2521 TGTGACTTTTGGACTGCGCCGAGACATCTCTGAGCGACTCCAGCTACGTCGTCAGGGCAAC 2580
Qy 2581 GCACGCTCCGCTGAAGTGGATGGCACCAGAGCTTATTTGAAGGATCTACACAATC 2640
Db 2581 GCACGCTCCGCTGAAGTGGATGGCACCAGAGCTTATTTGAAGGATCTACACAATC 2640
Qy 2641 AAGAGTGAAGTGTGCTTACGCGCATCTTCTCTGGAGATATTTTCACTGGGTGTGAAC 2700
Db 2641 AAGAGTGAAGTGTGCTTACGCGCATCTTCTCTGGAGATATTTTCACTGGGTGTGAAC 2700
Qy 2701 CCTTACCTCTGGCATCTCTGTCGACGCTAATCTTAAACTGATTCAGAGTGGATTTAAA 2760
Db 2701 CCTTACCTCTGGCATCTCTGTCGACGCTAATCTTCTAATAACTGATTCAGAGTGGATTTAAA 2760
Qy 2761 ATGGAGAGGCAATCTATGCGCAAGAAAGGATATATCTTTGTAATGCAATCTCTGCTGGCT 2820
Db 2761 ATGGAGAGGCAATCTATGCGCAAGAAAGGATATATCTTTGTAATGCAATCTCTGCTGGCT 2820
Qy 2821 TTTGACTCAAGGAAGCGGCGCATCTTCCCAACTGACTTCTATTTTAGGATGTGAGCTG 2880
Db 2821 TTTGACTCAAGGAAGCGGCGCATCTTCCCAACTGACTTCTATTTTAGGATGTGAGCTG 2880

Qy	181	CCATCATGTACCGAATGGTGCAGAGATCCCCAGAAAGCATCTTCAGTGTACCCCGAGGCGC	240
Db	232		291
		CCATCATCGTACCGAATGGTGCAGAGATCCCCAGAAAGCACCTCCAGTGTACCCCGAGGCGC	
Qy	241	CAGAGTGAAGGACCGGTATATGAACGCGCCACCGGTGAGAGTGGCCGAGTCTGGGTCCATC	300
Db	292		351
		CAGAGTGAAGGACCGGTATATGAACGCGCCACCGGTGAGAGTGGCCGAGTCTGGGTCCATC	
Qy	301	ACCCCTGCAAGTGCAGCTCGCCACCCACAGGGGACCTTTCCTGCCTCTGGGTCTTTAAAGCAC	360
Db	352		411
		ACCCCTGCAAGTGCAGCTCGCCACCCACAGGGGACCTTTCCTGCCTCTGGGTCTTTAAAGCAC	
Qy	361	AGCTCCCTGGGCTCCAGCGGCACTTTGATTTACAAACAGAGGAATCGTTTTCATGCCC	420
Db	412		471
		AGCTCCCTGGGCTCCAGCGGCACTTTGATTTACAAACAGAGGAATCGTTTTCATGCCC	
Qy	421	ATCTTGAACGTGACAGAGACCCAGGACGAGGAATACCTACTCCATATTTACAGAGCGAAAGC	480
Db	472		531
		ATCTTGAACGTGACAGAGACCCAGGACGAGGAATACCTACTCCATATTTACAGAGCGAAAGC	
Qy	481	GCCAACTACACAGTACTGTTTCACAGTGAATGTAAAGAGATACACAGCTGTATGTCTTAAGG	540
Db	532		591
		GCCAACTACACAGTACTGTTTCACAGTGAATGTAAAGAGATACACAGCTGTATGTCTTAAGG	
Qy	541	AGACCTTACTTTAGGAAGATGGAAAAACAGAGTGCATGTCTCTGGATCTCCGAGGGTGT	600
Db	592		651
		AGACCTTACTTTAGGAAGATGGAAAAACAGAGTGCATGTCTCTGGATCTCCGAGGGTGT	
Qy	601	CCGAGGCCACTGTGGAGTGGGTGCTCTGACGCTCCACAGGGAAGCTGTAAAGAGAA	660
Db	652		711
		CCGAGGCCACTGTGGAGTGGGTGCTCTGACGCTCCACAGGGAAGCTGTAAAGAGAA	
Qy	661	GGCCCTGCTGTTGTCAGAAAAAGGAAAAAGGTACTTTATGAGTTGTTTCGGAAACAGACATC	720
Db	712		771
		GGCCCTGCTGTTGTCAGAAAAAGGAAAAAGGTACTTTATGAGTTGTTTCGGAAACAGACATC	
Qy	721	AGATGCTGTGTAGAAATGCACCTGGGCGCGAAATGCACCAAGCTGTTTCACCATAGATCTA	780
Db	772		831
		AGATGCTGTGTAGAAATGCACCTGGGCGCGAAATGCACCAAGCTGTTTCACCATAGATCTA	
Qy	781	AACAGGCTCTCTAGAGCACACTGCCCCAGTTATTCCTGAAAGTGGGGGAAACCCCTGTGG	840
Db	832		891
		AACAGGCTCTCTAGAGCACACTGCCCCAGTTATTCCTGAAAGTGGGGGAAACCCCTGTGG	
Qy	841	ATCAGGTGTAAAGGCCATCCATGTGAAACCATGGAATTCGGGCTCACCTCGGAGCTGGAAGAC	900
Db	892		951
		ATCAGGTGTAAAGGCCATCCATGTGAAACCATGGAATTCGGGCTCACCTCGGAGCTGGAAGAC	
Qy	901	AAAGCCCTGGAGGAGGCGAGCTACTTTGAGATGAGTACCTTACCAAAACAGGACCATG	960
Db	952		1011
		AAAGCCCTGGAGGAGGCGAGCTACTTTGAGATGAGTACCTTACCAAAACAGGACCATG	
Qy	961	ATTCCGATCTCTTTGGCCTTTGTGCTCTCCGTGGGAAGGAACGACACCGGATATTACCC	1020
Db	1012		1071
		ATTCCGATCTCTTTGGCCTTTGTGCTCTCCGTGGGAAGGAACGACACCGGATATTACCC	
Qy	1021	TGCTCTTCTCTCAAAGCACCCAGCCAGTCAGCGTTGGTGAACCATCTCTAGAAAAAGGGTTT	1080
Db	1072		1131
		TGCTCTTCTCTCAAAGCACCCAGCCAGTCAGCGTTGGTGAACCATCTCTAGAAAAAGGGTTT	
Qy	1081	ATAAACGCTACCGCTCGCAAGAGAGATATGAATTCAGCCCGTACGAAAAAGTTCTGCTTC	1140
Db	1132		1191
		ATAAACGCTACCGCTCGCAAGAGAGATATGAATTCAGCCCGTACGAAAAAGTTCTGCTTC	
Qy	1141	TCAGTCAAGTTTAAAGGCTACCCAGAAATCCGATGCAAGTGGATCTTCTCTCAAGGCTCA	1200
Db	1192		1251
		TCAGTCAAGTTTAAAGGCTACCCAGAAATCCGATGCAAGTGGATCTTCTCTCAAGGCTCA	
Qy	1201	TTTCTTGTGAAACAGAGAGGCTCGAGGATGGGTACAGCATATCTAAATTTTCGGATCAT	1260
Db	1252		1311
		TTTCTTGTGAAACAGAGAGGCTCGAGGATGGGTACAGCATATCTAAATTTTCGGATCAT	
Qy	1261	AAGAAACAGCCAGGAGAGTACATATTTCTATGCAAGAAATGATACGCCCCAGTTTCACCAA	1320

Db	1312	 AAGAA CAAGC CAGAG AGTAC ATATCT ATGCA GAAAT ATGAT GACG CCAG GTCAC CAAA TGAT GACG CCAG GTCAC CAAA 1371
Qy	1321	 ATGTT CACGT GTAAT ATAAG AAGAA AACC TCAAG TGCT TAGCA AATG CCTC TAGC CAGC CAG CCAG 1380
Db	1372	 ATGTT CAGCT GAAAT ATAGA AAGAA AACC TCAAG TGCT TAGCA AATG CCTC TAGC CAGC CAG CCAG 1431
Qy	1381	 GGTCT GTGTT CCTCT GATGG CTAC CCGCT ACCC TCTTG GACCT GCGAA AGTG TTCGG AC 1440
Db	1432	 GGTCT GTGTT CCTCT GATGG CTAC CCGCT ACCC TCTTG GACCT GCGAA AGTG TTCGG AC 1491
Qy	1441	 AAATCT CCCAAT TTGCA CGGAG GAAAT TCCCA GAGAG GTTTC GGAAT ATAAA AGCTA AACGA 1500
Db	1492	 AAATCT CCCAAT TTGCA CGGAG GAAAT TCCCA GAGAG GTTTC GGAAT ATAAA AGCTA AACGA 1551
Qy	1501	 AAAGT GTTTGG CCAGT GGGTG TCGAG CAGTAC TCTAA ATATAG TGTAG TGAGC CGGAA AGGG 1560
Db	1552	 AAAGT GTTTGG CCAGT GGGTG TCGAG CAGTAC TCTAA ATATAG TGTAG TGAGC CGGAA AGGG 1611
Qy	1561	 CTTCT GGTCAA ATGCT GTGCG TACAA TTCTAT TGGGC ACGTC TTCG GAAAC CATCT TTTTTA 1620
Db	1612	 CTTCT GGTCAA ATGCT GTGCG TACAA TTCTAT TGGGC ACGTC TTCG GAAAC CATCT TTTTTA 1671
Qy	1621	 AACTCA CAGG CCCTT CCCTT TCAAT CACA AGA CAATC TCTCT TCTAT GCGAC CAATGG 1680
Db	1672	 AACTCA CAGG CCCTT CCCTT TCAAT CACA AGA CAATC TCTCT TCTAT GCGAC CAATGG 1731
Qy	1681	 CTCTGT CTCC CTTCA TTGTT GTTCT CATTT GTGTT GATCT GCACA CAATAC CAAAA AGCAA 1740
Db	1732	 CTCTGT CTCC CTTCA TTGTT GTTCT CATTT GTGTT GATCT GCACA CAATAC CAAAA AGCAA 1791
Qy	1741	 TTTAGT ACGAG AGTCAG CTGCAG ATGAT GCATC CAGGT GACTG GCCCT CCCTG GATAA CAGGTAC 1800
Db	1792	 TTTAGT ACGAG AGTCAG CTGCAG ATGAT GCATC CAGGT GACTG GCCCT CCCTG GATAA CAGGTAC 1851
Qy	1801	 TTTCTA CGTTGA CTTTCAG GGA CTATGA ATATGA CCTTAA GTGGG AGTTCC CGAGAG AGAAC 1860
Db	1852	 TTTCTA CGTTGA CTTTCAG GGA CTATGA ATATGA CCTTAA GTGGG AGTTCC CGAGAG AGAAC 1911
Qy	1861	 TTTAGA GTTTGG GAAGT CTCTGG GGTCTG GGGCTT TCGGG AGGGT GATGA ACCGCA CGGCC 1920
Db	1912	 TTTAGA GTTTGG GAAGT CTCTGG GGTCTG GGGCTT TCGGG AGGGT GATGA ACCGCA CGGCC 1971
Qy	1921	 TATGG CATTAG TAAAC CGGAG TCTCAA TTTCAG GTGGC GGTGA AGATG CTTAAA AGAGAAA 1980
Db	1972	 TATGG CATTAG TAAAC CGGAG TCTCAA TTTCAG GTGGC GGTGA AGATG CTTAAA AGAGAAA 2031
Qy	1981	 GCTGAC AGCTGT GAAAA AAGACT CTCAT GTGCG AGCTCA AAATG ATGAC CCAC CTGGGA 2040
Db	2032	 GCTGAC AGCTGT GAAAA AAGACT CTCAT GTGCG AGCTCA AAATG ATGAC CCAC CTGGGA 2091
Qy	2041	 CACCAT GACAA CATCGT GTAAT CTGCTG GGGCA TGCACA CTGT CAGGC CGCAGT GTACTT G 2100
Db	2092	 CACCAT GACAA CATCGT GTAAT CTGCTG GGGCA TGCACA CTGT CAGGC CGCAGT GTACTT G 2151
Qy	2101	 ATTTT TGAATA TTGTTG CTATG GTGAC CTCCT CAACT TACCTA AGAAG TAAAG AGAGA G 2160
Db	2152	 ATTTT TGAATA TTGTTG CTATG GTGAC CTCCT CAACT TACCTA AGAAG TAAAG AGAGA G 2211
Qy	2161	 TTTCA CAGGAC ATGGA CAGAG ATTTT TAAAG GAACAT PAA TTTCAG TTCTT TACCTT CTTTC 2220
Db	2212	 TTTCA CAGGAC ATGGA CAGAG ATTTT TAAAG GAACAT PAA TTTCAG TTCTT TACCTT CTTTC 2271
Qy	2221	 CAGGCA CATTC CAAA TTCAG CATG CCTGG TTCACG AGAAGT TCAGT TACAC CGCC CTTG 2280
Db	2272	 CAGGCA CATTC CAAA TTCAG CATG CCTGG TTCACG AGAAGT TCAGT TACAC CGCC CTTG 2331
Qy	2281	 GATCAG CTCTC AGGGT TCAA TGGGA ATTCAA TTTCTT CTGGA AGTGA GATTG AATGAA 2340
Db	2332	 GATCAG CTCTC AGGGT TCAA TGGGA ATTCAA TTTCTT CTGGA AGTGA GATTG AATGAA 2391
Qy	2341	 AACCA AAGAG CGCTG CGCAG AAGA GAGG AGAGAT TTGAA CCTGCT GCGAT TTTGA AGAC 2400

Db 2392 AACACAGAGGCTGGCAGAGAGAGGAGGAAGATTGAACTGCTGACGTTTGAAGAC 2451
Qy 2401 CTCCTTTGCTTTGCTTACCAAGTGGCCAAAGGATGGAATTCCTGAGTTCAAGTCGTGT 2460
Db 2452 CTCCTTTGCTTTGCTTACCAAGTGGCCAAAGGATGGAATTCCTGAGTTTCAAGTCGTGT 2511
Qy 2461 GTCCACAGAGACTGGCAGCAGCAAGATGCTGCTACCCACCGGAGGTGGTGAAGATC 2520
Db 2512 GTCCACAGAGACTGGCAGCAGCAAGATGCTGCTACCCACCGGAGGTGGTGAAGATC 2571
Qy 2521 GTGACTTTGGACTGGCCCGGAGACATCTCTGAGCGACTCCAGCTACGTCGTGAGGGGCAAC 2580
Db 2572 GTGACTTTGGACTGGCCCGGAGACATCTCTGAGCGACTCCAGCTACGTCGTGAGGGGCAAC 2631
Qy 2581 GCACGCTGGCTGAGTGGATGGACCGGAGAGCTTATTGAAGGATCTACAAATC 2640
Db 2632 GCACGCTGGCTGAGTGGATGGACCGGAGAGCTTATTGAAGGATCTACAAATC 2691
Qy 2641 AAGAGTGAGCTGCTGCTACGCGCATCTTCTCTGGGAGATATTTTCACTGGGTGTAAC 2700
Db 2692 AAGAGTGAGCTGCTGCTACGCGCATCTTCTCTGGGAGATATTTTCACTGGGTGTAAC 2751
Qy 2701 CTTTACCTTGGCATCTCTGCTGAGCGCTAACTTCTATAAAGTGAATTCAGAGTGAATTTAAA 2760
Db 2752 CTTTACCTTGGCATCTCTGCTGAGCGCTAACTTCTATAAAGTGAATTCAGAGTGAATTTAAA 2811
Qy 2761 ATGGAGCAGCATCTATGCCACAGAGGATATATCTTTGAATGAATCTCTGCTGGGT 2820
Db 2812 ATGGAGCAGCATCTATGCCACAGAGGATATATCTTTGAATGAATCTCTGCTGGGT 2871
Qy 2821 TTTGACTCAAGAGGCGGCATCTTCTCCACCTGACTTTTATAGGATGTCAGCTG 2880
Db 2872 TTTGACTCAAGAGGCGGCATCTTCTCCACCTGACTTTTATAGGATGTCAGCTG 2931
Qy 2881 GCAGAGGCGAAGAGAGC-----ATGTATCAGAAATCATCATCC 2917
Db 2932 GCAGAGGCGAAGAGAGAGATGATCAGAACTGGGTGGCAAGCTCCAGAAATCATCATCC 2991
Qy 2918 ATCTACCAAAACAGGCGGCTCTAGCAGAGAGGCGGCTCAGAGCC-CAGTGGCCACAG 2976
Db 2992 ATCTACCAAAACAGGCGGCTCTAGCAGAGAGGCGGCTCAGAGCGCGCATCGCCACAG 3051
Qy 2977 CGCAGGTGAAGATTCACAGAGAAAGTGTAGCAGAGGCGCTTGGACCCGCCACCT 3036
Db 3052 GCCCAGGTGAAGATTCACAGAGAAAGTGTAGCAGAGGCGCTTGGACCCGCCACCT 3111
Qy 3037 AGCAGGCTGTAGACCGCAGAGCAAGATTAGCTCTGAGGAGGCGCTCAGAGCGCTCAGCG 3096
Db 3112 AGCAGGCTGTAGACCGCAGAGCAAGATTAGCTCTGAGGAGGCGCTCAGAGCGCTCAGCG 3171
Qy 3097 CGTTCCTGCTGGACTTTTCTAGATGCTGCTGCGCAATTAATTAAGTGAATCTTAT 3156
Db 3172 CGTTCCTGCTGGACTTTTCTAGATGCTGCTGCGCAATTAATTAAGTGAATCTTAT 3231
Qy 3157 AAATCAAACTCTCTCGCAGCGGAGAGCAATTAATTAAGTGAATCTTAT 3216
Db 3232 AAATCAAACTCTCTCGCAGCGGAGAGCAATTAATTAAGTGAATCTTAT 3291
Qy 3217 GCCTACCTTGGGGGCTTT-----CCACGAGCTTGGAGGGAAGCCATCTATCTGAA 3268
Db 3292 GCCTACCTTGGGGGCTTTCCAGGCCCCCAGGCTTGGAGGGAAGCCATCTATCTGAA 3351
Qy 3269 ATATAGTATATCTTGTAAATACGTGAAACAAACCAACCCCGTTTTTGTCTAAGGGAAG 3328
Db 3352 ATATAGTATATCTTGTAAATACGTGAAACAAACCAACCCCGTTTTTGTCTAAGGGAAG 3411
Qy 3329 CTAATATGATTTTAAATCTATGTTTAAATACGTGAAATCTTATCTATCTATTTA 3389
Db 3412 CTAATATGATTTTAAATCTATGTTTAAATACGTGAAATCTTATCTATCTATTTA 3471
Qy 3389 GTGATATATTTATGATGGAATAAATCTTCTCTGTAATAAAAAA 3438
Db 3472 GTGATATATTTATGATGGAATAAATCTTCTCTGTAATAAAAAA 3521

RESULT 12

US-08-434-878-1
; Sequence 1, Application US/08434878
; Patent No. 5957865
; GENERAL INFORMATION:
; APPLICANT: Bennett, Brian D.
; APPLICANT: Broz, Susan D.
; APPLICANT: Matthews, William
; APPLICANT: Zeigler, Francis C.
; TITLE OF INVENTION: AGONIST ANTIBODIES AGAINST THE FLK2/FLT3 RECEPTOR AND USES THE
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/434,878
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Hasak, Janet E.
; REGISTRATION NUMBER: 28,616
; REFERENCE/DOCKET NUMBER: 879
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-1896
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3521 bases
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-434-878-1

Query Match 96.8%; Score 3344; DB 2; Length 3521;
Best Local Similarity 98.5%; Pred. No. 0;
Matches 3418; Conservative 0; Mismatches 20; Indels 32; Gaps 3;

Qy 1 CGGCGCTGGCTACCGCGGCTCCGAGGCGCATCGGGGCTTGGCGAGCGGCGGACCGG 60
Db 52 CGGCGCTGGCTACCGCGGCTCCGAGGCGCATCGGGGCTTGGCGAGCGGCGGACCGG 111
Qy 61 CGGCTGCTGCTGCTGTTGTTTGTGCTAGTATGTTCTTGAGACCGTTACAAACCAAGAC 120
Db 112 CGGCTGCTGCTGCTGTTGTTTGTGCTAGTATGTTCTTGAGACCGTTACAAACCAAGAC 171
Qy 121 CTGCTGTGATCAAGTGTGTTTAAATCAGTCATGAGAAACAATGGCTCATCAGCGGGAAG 180
Db 172 CTGCTGTGATCAAGTGTGTTTAAATCAGTCATGAGAAACAATGGCTCATCAGCGGGAAG 231
Qy 181 CCATCATCTGTAAGTGTGCGAGGATCCCGAGAGACCTCCAGAGTGTACCCCGAGGCGC 240
Db 232 CCATCATCTGTAAGTGTGCGAGGATCCCGAGAGACCTCCAGAGTGTACCCCGAGGCGC 291
Qy 241 CAGAGTGAAGGACGCTATATGAAGCGCCACCGTGGAGGTGGCGGAGTCTGGGTCCATC 300
Db 292 CAGAGTGAAGGACGCTATATGAAGCGCCACCGTGGAGGTGGCGGAGTCTGGGTCCATC 351
Qy 301 ACCCTGCAAGTGCAGCTCGCCACCCAGGGGACCTTTCTCTGCTCTGGGTCTTTAAGCAC 360

Db 532 GCCAACTACACAGTACTGTTTCACAGTGAATGTAAGAGATACACAGCTGTATGTGCTAAGG 591
Qy 541 AGACCTTACTTTAGGAAGATGAAAAACAGATGCACTGCTCTGCAATCTCCGAGGGTGT 600
Db 592 AGACCTTACTTTAGGAAGATGAAAAACAGATGCACTGCTCTGCAATCTCCGAGGGTGT 651
Qy 601 CCGGAGCCCACTGTGGAGTGGTGTCTGCAAGCTCCCAAGGAAAGCTGTAAGAAGAA 660
Db 652 CCGGAGCCCACTGTGGAGTGGTGTCTGCAAGCTCCCAAGGAAAGCTGTAAGAAGAA 711
Qy 661 GGCCCTGCTGTGTGTCAGAAAGGAGAAAGGTACTTTCAATGAGTTGTTTCGGAACAGACATC 720
Db 712 GGCCCTGCTGTGTGTCAGAAAGGAGAAAGGTACTTTCAATGAGTTGTTTCGGAACAGACATC 771
Qy 721 AGATGCTGTGTAGAAATGCACTGGGCGCGAATGCAACCAAGCTGTTCACCATGATCTA 780
Db 772 AGATGCTGTGTAGAAATGCACTGGGCGCGAATGCAACCAAGCTGTTCACCATGATCTA 831
Qy 781 AACCAAGCTCTCAGAGACACTGCCCGAGTTATTCCTGAAAGTGGGGGAAACCTTGTGG 840
Db 832 AACCAAGCTCTCAGAGACACTGCCCGAGTTATTCCTGAAAGTGGGGGAAACCTTGTGG 891
Qy 841 ATCAAGGTGTAAAGCCATCTCAATGCAACCATGATTCGGGCTCACCTGGGAGCTGGAAGAC 900
Db 892 ATCAAGGTGTAAAGCCATCTCAATGCAACCATGATTCGGGCTCACCTGGGAGCTGGAAGAC 951
Qy 901 AAAGCCCTGGAGGAGGAGCTACTTTGAGATGAGTACCTACTCCACAAACAGGACCATG 960
Db 952 AAAGCCCTGGAGGAGGAGCTACTTTGAGATGAGTACCTACTCCACAAACAGGACCATG 1011
Qy 961 ATTCCGATTCCTTTGGCCTTTGTTGTTCTCCGTTGGGAGGAAACGACACCGGATATTACCC 1020
Db 1012 ATTCCGATTCCTTTGGCCTTTGTTGTTCTCCGTTGGGAGGAAACGACACCGGATATTACCC 1071
Qy 1021 TGCTCTTCTCAAAGCAGCCAGCCAGTCAGCGTTGGTGACCATCTTAGAAAGAGGTTT 1080
Db 1072 TGCTCTTCTCAAAGCAGCCAGCCAGTCAGCGTTGGTGACCATCTTAGAAAGAGGTTT 1131
Qy 1081 ATAAAGCTACAGCTCGAAGAGAGATGAAATTTGACCGGTACGAAAGTTCTGCTTC 1140
Db 1132 ATAAAGCTACAGCTCGAAGAGAGATGAAATTTGACCGGTACGAAAGTTCTGCTTC 1191
Qy 1141 TCAGTCAGGTTTAAAGGATACCAAGATCCGATGCGAGTGAGTCTCTCAAGCCCTCA 1200
Db 1192 TCAGTCAGGTTTAAAGGATACCAAGATCCGATGCGAGTGAGTCTCTCAAGCCCTCA 1251
Qy 1201 TTTCTTGTGAACAGAGAGGCTCGAGATGGGTACAGCATATCTAAATTTTGGATCAT 1260
Db 1252 TTTCTTGTGAACAGAGAGGCTCGAGATGGGTACAGCATATCTAAATTTTGGATCAT 1311
Qy 1261 AAGAACAGCCAGGAGAGTACATATTTCTATGCAAGAAATGATGACGCCAGTTTCAACAAA 1320
Db 1312 AAGAACAGCCAGGAGAGTACATATTTCTATGCAAGAAATGATGACGCCAGTTTCAACAAA 1371
Qy 1321 ATGTTACGCTGAATATAAGAAAGAAACCTCAAGTGTAGCAATGCTCAGCAGCCAG 1380
Db 1372 ATGTTACGCTGAATATAAGAAAGAAACCTCAAGTGTAGCAATGCTCAGCAGCCAG 1431
Qy 1381 GCGTCTGTTCTCTGATGGTACCGCTACCTCTTGGACCTGGAAGAGTGTTCGGAC 1440
Db 1432 GCGTCTGTTCTCTGATGGTACCGCTACCTCTTGGACCTGGAAGAGTGTTCGGAC 1491
Qy 1441 AAATCTCCCAATTCGACGGAGGAAATCCCAAGAGGATTTTGGAAATAAAGGCTTAACAGA 1500
Db 1492 AAATCTCCCAATTCGACGGAGGAAATCCCAAGAGGATTTTGGAAATAAAGGCTTAACAGA 1551
Qy 1501 AAAGTGTGTCAGTGGGTGTCGAGCAGTACTCTAAATATGAGTGGGCGGGAAAGGG 1560
Db 1552 AAAGTGTGTCAGTGGGTGTCGAGCAGTACTCTAAATATGAGTGGGCGGGAAAGGG 1611
Qy 1561 CTTCTGCTCAATGCTGTGGTACAAATCTATGGGCGAGCTTTCGGAACCATCTTTTAA 1620
Db 1612 CTTCTGCTCAATGCTGTGGTACAAATCTATGGGCGAGCTTTCGGAACCATCTTTTAA 1671

Qy 1621 AACTCACAGGCCCTTCCCTTTTCATCCAGAGCAACATCTCTTCTATGCGACCAATTGGG 1680
Db 1672 AACTCACAGGCCCTTCCCTTTTCATCCAGAGCAACATCTCTTCTATGCGACCAATTGGG 1731
Qy 1681 CTCTGTCTCCCTTTCATTTGTTTCTCATTTGTTGATCTGCCCAAAATACAAAAGCAA 1740
Db 1732 CTCTGTCTCCCTTTCATTTGTTTCTCATTTGTTGATCTGCCCAAAATACAAAAGCAA 1791
Qy 1741 TTTAGGTACGAGAGTCAGCTGACAGTATCCAGTGTAGTGGCCCCCTCGATAACAGATAC 1800
Db 1792 TTTAGGTACGAGAGTCAGCTGACAGTATCCAGTGTAGTGGCCCCCTCGATAACAGATAC 1851
Qy 1801 TTTACGTTGATCTCAGGAGCTATGAAATATGACCTTAAGTGGGAGTTCCTCGAGAGAGAAC 1860
Db 1852 TTTACGTTGATCTCAGGAGCTATGAAATATGACCTTAAGTGGGAGTTCCTCGAGAGAGAAC 1911
Qy 1861 TTAGAGTTTGGGAAGGCTCTGGGCTCTGGGCTCTGGGAGGGTGAAGAACGCCACGGCC 1920
Db 1912 TTAGAGTTTGGGAAGGCTCTGGGCTCTGGGCTCTGGGAGGGTGAAGAACGCCACGGCC 1971
Qy 1921 TATGGCATTAAGTAAACGGGAGTCTCAATTCAGGTGGCGGTGAAGATGCTAAAGAGAAA 1980
Db 1972 TATGGCATTAAGTAAACGGGAGTCTCAATTCAGGTGGCGGTGAAGATGCTAAAGAGAAA 2031
Qy 1981 GCTGACAGCTGTGAAAGAGAGCTCTCATGTGCGAGCTCAAAATGATGACCCACCTGGGA 2040
Db 2032 GCTGACAGCTGTGAAAGAGAGCTCTCATGTGCGAGCTCAAAATGATGACCCACCTGGGA 2091
Qy 2041 CACCATGACATCGTGAATCTCTGCGGGCATGCACTGTGAGGGCCAGTGTACTTG 2100
Db 2092 CACCATGACATCGTGAATCTCTGCGGGCATGCACTGTGAGGGCCAGTGTACTTG 2151
Qy 2101 ATTTTCAATATTTCTGCTATGTTGACCTCTCAACTACCTAAAGAGTAAAGAGAGAG 2160
Db 2152 ATTTTCAATATTTCTGCTATGTTGACCTCTCAACTACCTAAAGAGTAAAGAGAGAG 2211
Qy 2161 TTTTCAAGGACATGACAGAGATTTTAAAGGAAATATTTTCAAGTCTTACCCTACTTTC 2220
Db 2212 TTTTCAAGGACATGACAGAGATTTTAAAGGAAATATTTTCAAGTCTTACCCTACTTTC 2271
Qy 2221 CAGGACATTAATTTCCAGCATGCTGTTTACAGAGATGTTCAAGTCTTACCCTGCTT 2280
Db 2272 CAGGACATTAATTTCCAGCATGCTGTTTACAGAGATGTTCAAGTCTTACCCTGCTT 2331
Qy 2281 GATCAGCTCTCAGGTTCAATGGGAATCAATTCATCTGAAGATGAGATTGAATATCAA 2340
Db 2332 GATCAGCTCTCAGGTTCAATGGGAATCAATTCATCTGAAGATGAGATTGAATATCAA 2391
Qy 2341 AACCAAGAGGCTGGCAGAGAGAGAGAGATTTGAACGCTGCTGACGTTTGAAGAC 2400
Db 2392 AACCAAGAGGCTGGCAGAGAGAGAGAGAGATTTGAACGCTGCTGACGTTTGAAGAC 2451
Qy 2401 CTCCTTTGCTTTGCTATCCAGTGGCCAAAGGATGAAATTTCTGGAGTTCAAGTCTGT 2460
Db 2452 CTCCTTTGCTTTGCTATCCAGTGGCCAAAGGATGAAATTTCTGGAGTTCAAGTCTGT 2511
Qy 2461 GTCCACAGACCTGGCAGCAGAGATGTTGTTGTCACCCACGGAAGGTTGTTGAATC 2520
Db 2512 GTCCACAGACCTGGCAGCAGAGATGTTGTTGTCACCCACGGAAGGTTGTTGAATC 2571
Qy 2521 TGTGACTTTGACTGGCCCGAGACATCTGAGCGACTCCAGTACGCTGCTCAGGGGCAAC 2580
Db 2572 TGTGACTTTGACTGGCCCGAGACATCTGAGCGACTCCAGTACGCTGCTCAGGGGCAAC 2631
Qy 2581 GCACGGCTGCGGTGAAGTGGATGGCACCCGAGAGCTTATTTGAAGGGATCTACACATC 2640
Db 2632 GCACGGCTGCGGTGAAGTGGATGGCACCCGAGAGCTTATTTGAAGGGATCTACACATC 2691
Qy 2641 AAGATGACGCTGTGCTCCTACGGCATCCCTTCTCTGGAGATATTTTCACTGGGTGTGAAC 2700
Db 2692 AAGATGACGCTGTGCTCCTACGGCATCCCTTCTCTGGAGATATTTTCACTGGGTGTGAAC 2751

2701 CCTTACCTGGCATTCTCTGACGCTAACTTCTATAAACTGATTGAGAGTGATTTAA 2760
2752 CCTTACCTGGCATTCTCTGACGCTAACTTCTATAAACTGATTGAGAGTGATTTAA 2811
2761 ATGGAGAGCCATTCTATGCCAGAGAGGATATCTTTGTAATGCAATCTCTGGGCT 2820
2812 ATGGAGAGCCATTCTATGCCAGAGAGGATATCTTTGTAATGCAATCTCTGGGCT 2871
2821 TTTGACTCAAGAGAGGCGCATCTTCCCAACCTGACTTCTATTTTAGATGTGAGCTG 2880
2872 TTTGACTCAAGAGAGGCGCATCTTCCCAACCTGACTTCTATTTTAGATGTGAGCTG 2931
2881 GCAGAGGAGAGAGGCGCATCTTCCCAACCTGACTTCTATTTTAGATGTGAGCTG 2917
2932 GCAGAGGAGAGAGGCGCATCTTCCCAACCTGACTTCTATTTTAGATGTGAGCTG 2991
2918 ATCTACCAAAACAGCGCGCCCTCTAGCAGAGAGGCGGCTCAGAGCC-CAGTCGCGACAG 2976
2992 ATCTACCAAAACAGCGCGCCCTCTAGCAGAGAGGCGGCTCAGAGCCCGCATCGCCACAG 3051
2977 CGCAGGTGAAGATTCAAGAGAGAGAGGTTAGCAGAGAGGCGCTTGGACCCCGCCACCT 3036
3052 GCCCAGGTGAAGATTCAAGAGAGAGAGGTTAGCAGAGAGGCGCTTGGACCCCGCCACCT 3111
3037 AGCAGGCTGTAGACGCGCAGAGATTAGCTCGCTCTGAGAGAGGCGCTTGGACCCCGCCACCT 3096
3112 AGCAGGCTGTAGACGCGCAGAGATTAGCTCGCTCTGAGAGAGGCGCTTGGACCCCGCCACCT 3171
3097 CGTTGCTTGGCTGAGCTTTTCTCTAGATGCTGTCTGCCATTCTCAAAAGTACTTCTAT 3156
3172 CGTTGCTTGGCTGAGCTTTTCTCTAGATGCTGTCTGCCATTCTCAAAAGTACTTCTAT 3231
3157 AAATCAAACTCTCTCGACAGCGGAGAGCAATATGAGACTTGTGTTGGAGGCC 3216
3232 AAATCAAACTCTCTCGACAGCGGAGAGCAATATGAGACTTGTGTTGGAGGCC 3291
3217 GCCTACCTCTGGGCGCTT-----CCAGAGCTTGGAGGAGAGGCAATGATCTGAA 3268
3292 GCCTACCTCTGGGCGCTTCCAGCGCCCGCCAGGCTTGGAGGAGAGGCAATGATCTGAA 3351
3269 ATATGATATATTTCTGTAATATCTGTAATATCTGTAATATCTGTAATATCTGTAAT 3328
3352 ATATGATATATTTCTGTAATATCTGTAATATCTGTAATATCTGTAATATCTGTAAT 3411
3329 CTAATATGATATTTTAAATATCTGTAATATCTGTAATATCTGTAATATCTGTAAT 3388
3412 CTAATATGATATTTTAAATATCTGTAATATCTGTAATATCTGTAATATCTGTAAT 3471
3389 GTGATATATTTTATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3438
3472 GTGATATATTTTATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3521

RESULT 14

US-07-977-451-3
; Sequence 3, Application US/07977451
; Patent No. 5270458
; GENERAL INFORMATION:
; APPLICANT: Lemischka, Ibor R.
; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL
; TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ImClone Systems Incorporated
; STREET: 180 Varick Street
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10014
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/977,451

FILING DATE: 19921119

CLASSIFICATION: 536

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US UNASSIGNED

FILING DATE: 12-NOV-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/906,397

FILING DATE: 26-JUN-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US PCT/US92/05401

FILING DATE: 26-JUN-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: TW 81102961

FILING DATE: 15-APR-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US PCT/US92/02750

FILING DATE: 02-APR-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/813,593

FILING DATE: 24-DEC-1991

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/793,065

FILING DATE: 15-NOV-1991

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/728,913

FILING DATE: 28-JUN-1991

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/679,666

FILING DATE: 02-APR-1991

ATTORNEY/AGENT INFORMATION:

NAME: Peit, Irving N.

REGISTRATION NUMBER: 28,601

REFERENCE/DOCKET NUMBER: LEM-3-7P

TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-645-1405

TELEFAX: 212-645-2054

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 3501 base pairs

TYPE: NUCLEIC ACID

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: CDNA

HYPOTHETICAL: NO

ANTI-SENSE: NO

FRAGMENT TYPE: N-terminal

FEATURE:

NAME/KEY: CDS

LOCATION: 58..3039

FEATURE:

NAME/KEY: mat_peptide

LOCATION: 139..3036

FEATURE:

NAME/KEY: sig_peptide

LOCATION: 58..138

US-07-977-451-3

Query Match: 64.5%; Score 2227.2; DB 1; Length 3501;

Best Local Similarity 79.6%; Pred. No. 0;

Matches 2775; Conservative 0; Mismatches 653; Indels 56; Gaps 10;

QY 8 GGCTACCGCGCTCCGAGGCGATCGCGGCGTGGCGCAGCGCAGCGCGCGCTGC 67

DB 35 GGGGACCCCGGCTCCGAGGCGCATGCCGCGTGGC---GCGCGCGCGGCGCGTGC 91

QY 68 TGCTGCTGTTGTTTTTGTGTCAGTAATGATTTTGTAGACCGTTACAAACCAAGACCTGCTG 127

DB 92 CGCTGCTGTTGTTTTTGTGTCAGTAATGATTTTGTAGACCGTTACAAACCAAGACCTGCTG 151

QY 128 TGATCAAGTGTGTTTTTAATCAAGTCATGAGAACATGAGTGGCTCATCAGCGGGAAGCCATCAT 187

152 TGATCAAGTGTGTTTAAATCAATCAATAAGAACTAATGATTCATCAGTGGGAAAGTCATCAT 211
188 CGTACCGAATGTCGCGAGGATCCCAAGAGAGCTCCAGTGTACCCGAGGCGCCAGAGTG 247
212 CATATCCCATGATATCAGAAATCCCGGAAGAGCTCGGGTGTGGTGTGAGACCCCAAGAGCT 271
248 AAGGACGGGTATATGAAGCGGCACCGTGGAGGTGGCCGAGTCTGGGTCCATCACCCTCC 307
272 CAGGCACAGTGTACGAAGCTGCGGTGTGGAAGTGGATGTATCTGCTTCATCACAATCC 331
308 AAGTCAGCTGCCACCCAGGGGACCTTTCCTGCTCTGGGTCTTAAAGCAGCAGCTCCC 367
332 AAGTCTGCTGATGCGATGCGCCAGGAACTTTCCTGCTCTGGGTCTTAAAGCAGCAGCTCCC 391
368 TGGGTCGCGAGCGGACATTTGATTTTACAAAACAGAGGAATCGTTTCCATGGCCATCTTGA 427
392 TGAATTGGCCAGCCATTTTGAATTTACAAAACAGAGGAGTGTGTTTCCATGGTCAATTTGA 451
428 ACGTCACAGAGACCCAGGCGAGGAATACCTACTCCATATTCAGAGCGAAACGGCCAACT 487
452 AAATGACAGAAACCCAGCTGGAGGAATACCTACTTTTATTCAGAGTGAAGCTACCAATT 511
488 ACACAGTACTGTTACAGTGAATGTAAGAGATACACAGCTGTATGTCTTAAGGAGACCTT 547
512 ACACAATATTTGTTACAGTGAATGTAAGAAATACCTGCTTTACACATTTAAGAGACCTT 571
548 ACTTTAGAAAGATGAAAACAGAGATGCACTGCTCTGCAATCTCCGAGGGTGTCCGAGC 607
572 ACTTTAGAAAATGAAAACAGAGACGCGCTGCTGCAATCTGAGAGCGTTCAGAGC 631
608 CCAGTGTGGAGTGGGTGCTCTCAGCTCCACAGGGAAGCTGTAAGAGAAAGGCCCTG 667
632 CGATCGTGAATGGGTGCTTTGCGAATTCACAGGGGAAAGCTGTAAAGAAAGAAATGCCAG 691
668 CTGTTGTACAGAAAGAGGAAAGTACTTCATGAGTTGTTCCGAAACAGACATCAGATGCT 727
692 CTGTTGTTAAAAGAGGAAAGTGTCTTCATGAATTTATTTGGAGCGACATTAAGTGTCT 751
728 GTGCTAGAAATGCACTGGGCGCGAATGCAACAAGCTGTTTCAACATAGATCTTAAATCAA 811
752 GTGCCAGAAATGAATCGGCGAGGAAATGCACAGGCTGTTTCAACATAGATCTTAAATCAA 811
788 CTCCTCAGAGCACACTGCGCCAGTATTTCTGAAAGTGGGGAACCTTGTGGATCAGGT 847
812 CTCCTCAGACACACTGCGCAATTTATTTCTTAAAGTAGGGGAACCTTTATGGATTAAGGT 871
848 GTAAGGCGCATCATGTGAACATGGATTTCGGGCTCACCTGGGAGCTGGAAGACAAAGCCC 907
872 GCAAGCTGTTGATGTGAACCATGGATTTCGGGCTCACCTGGGAATTTAGAAAACAAGCAC 931
908 TGGAGGAGGCGAGCTACTTTGAGATGAGTACTCTCAACAAACAGGACCATGATTCGGA 967
932 TCGAGGAGGCGCAACTACTTTGAGATGAGTACTCTTCAACAAACAGAACTATGATACGGA 991
968 TTCTCTGGCCCTTTGTGTTCTCGGGAAGAGAACGACACCGGATATTAACCTGCTCTT 1027
992 TTCTGTTTGTCTTTGTATTCATCAGTGGCAAGAAACGACACCGGATATTAACCTGCTCT 1051
1028 CCTCAAGCACCCAGGCGAGTGGTGGTGAACCATCTTAGAAAAGGGTTTATAAAGC 1087
1052 CTTCAAGCATCCCATGTCATCAGCTTTGGTTTGGTTTACCATCTGAGAAAGGGAATTTAAAGT 1111
1088 CTACCAGCTCGAAGAGAGATGAAATTTGACCCGTACGAAAAGTTCTGCTTCTCAGTCA 1147
1112 CTACCAATTTCAAGTGAAGATTTGAAATTTGACCAATATGAAGAGTTTGTGTTTCTGTC 1171
1148 GGTTTAAAGCGTACCCAGAAATCCGATGCAAGTGGATCTTCTCAAGCCCTCAATTCCTT 1207
1172 GGTTTAAAGCGTACCCCAAAATCAGATGTAGTGGACCTTCTCTCGAAAATCAATTCCTT 1231
1208 GTGAACAGAGGCGCTGGAGATGGGTACAGCATATCTAAATTTTGGCATCATAGAACA 1267

1232 GTGAGCAAAAGGGTCTTTGATAACGGATACAGCATATCTCAAGTTTGTCAATCATAAAGCAC 1291
1268 AGCCAGGAGTACATATTTCTATGAGAAAATGATGACGCCCAAGTTTACCAAAATGTTCA 1327
1292 AGCCAGGAGAAATATATATTTCCATGCAAAAATGATGATGCCAAATTTACCAAAATGTTCA 1351
1328 CGCTGAATATAAGAAAAGAACTCAAGTGTAGCAAAATGCTCAGCCAGCAGCGCTCT 1387
1352 CGCTGATATAAGAAAGGAACTCAAGTGTGCGAGAGCATCGGCAAGTCAAGCGCTCT 1411
1388 GTTCTCTGATGGCTACCCGCTACCCCTTTGGAACCTTGGAGAAAGTGTTCGACAAAATCTC 1447
1412 GTTCTCTGGATGATACCCATTAACCATCTTGACCTTGGAAAGTGTTCAGACAAGTCTC 1471
1448 CCAATTTGACGAGGAAATCCAGAAAGGAGTTTGGAAATAAAAGGCTTAACAGAAAAGTGT 1507
1472 CCACTGACAGAGAGATACAGAAAGGAGTCTGGAATAGAAGGCTTAACAGAAAAGTGT 1531
1508 TTGGCCAGTGGGTGTCGAGCAGTACTCTAAATATAGTACGAGCGCGGAAAGGCTTCTGG 1567
1532 TTGGACAGTGGGTGTCGAGCAGTACTCTAAACATGAGTGAAGCCATAAAGGGGTTCCTGG 1591
1568 TCAAAATGCTGTGCTGTAACAATTTATGGGCACTGCTTTCGAAAACCATCTTTTAAATCTAC 1627
1592 TCAAGTGTGTCATACAATTCCTTGGCACATCTTGTGAGAGCATCTTTTAAATCTCTC 1651
1628 CAGGCGCTTCCCTTTTCATCAAGACAACATCTCTTCTATGCGACATTTGGGCTCTGTC 1687
1652 CAGGCGCTTCCCTTTTCATCAAGACAACATCTCTTCTATGCAACAAATGGTGTGTC 1711
1688 TCCCTTTCATTTGTTTCTCATTTGTTGATCTGCCACAATAACAAAAGCAATTTAGGT 1747
1712 TCCTTCTCATTTGCTGTTTAACTGCTAAATTTGTCAAGTACAAAAGCAATTTAGGT 1771
1748 ACGAGAGTCAGCTGCAGATGATCCAGGTGACTTGGCCCTCGATAACGAGTACTTCTACG 1807
1772 ATGAAGCCAGCTACAGATGTTGAGGTGACCGGCTCTCTCAGATATAGTACTTCTACG 1831
1808 TTGACTTCAGGAGCTATGAATATGACCTTAACTGGAGTTCCTCGAGAGAGAACTTAGAGT 1867
1832 TTGATTTCAAGAAATATGAATATGATCTCAATGGAGTTCCTCAAGAGAAAATTTAGAGT 1891
1868 TTGGAAAGGTCTGGGCTGCGGCTTTCGGAGAGGTGATGAACGCCACGGCTATGGCA 1927
1892 TTGGAAAGGTACTAGGATCAGTGTCTTTTGGAAAAGTGTGAACGCCACAGCTTATGGAA 1951
1928 TTAGTAAAACGGAGTCTCAATTCAGGTGGCGGTGAGATGCTTAAAGAGAAAGCTGACA 1987
1952 TTAGCAAAACAGAGGTCTCAATTCAGGTTCGCTCAAAATGCTGAAAAGAAAAGCAGACA 2011
1988 GCTGTAAAAGAAAGCTCTCATGTGCGAGCTCAAAATGATGACCCACCTGGGACACCATG 2047
2012 GCTGTAAAAGAGAGCACTCATGTGAGAACTCAAGATGATGACCCAGCTGGGAAGCCAG 2071
2048 ACAACATGTAATCTGCTGGGGCATGCACTGTCAAGGCGCAGTGTACTTGAATTTTGG 2107
2072 AGAATATTTGAAACCTGCTGGGGCGTGCACACTGTCAAGGACCAATTTACTTGAATTTG 2131
2108 AATATTTGTTGATGTCGCTCCTCACTACCTAAGAAAGTAAAGAGAGAGTTCACA 2167
2132 AATATGTTGCTATGTTGATCTCTCAACTATCTAAGAAAGTAAAGAGAGAAATTTTCA 2191
2168 GGACATGACAGAGATTTTAAAGAAACATAATTTTCAAGTTCTTACCTACTTTCAGGAC 2227
2192 GGACTTGGACAGAGATTTTCAAGGAACATAATTTTCAAGTTTTCACCTTTTCCATCTAC 2251
2228 ATTCAAAATTCAGCATGCTGTTTCAAGAGAGTTCAGTTACACCCGCTTTGGATCAGC 2287
2252 ATCAAAATTCAGCATGCTGTTTCAAGAGAGTTCAGATACACCCGACTCGGATCAAA 2311
2288 TCTCAGGGTTCATTTGGGAATTCATTTCTTGAAGATGAGATTTGAATATGAAAACCGA 2347
2312 TCTCAGGGCTTCATGGGAATTCATTTCTCTGAGATGAAATTTGAATATGAAAACCGA 2371


```

,
, TOPOLOGY: linear
, MOLECULE TYPE: cDNA
, HYPOTHETICAL: NO
, ANTI-SENSE: NO
, FRAGMENT TYPE: N-terminal
, FEATURE:
,   NAME/KEY: CDS
,   LOCATION: 58..3039
,   FEATURE:
,   NAME/KEY: mat_peptide
,   LOCATION: 139..3036
,   FEATURE:
,   NAME/KEY: sig_peptide
,   LOCATION: 58..138
, US-08-252-517-3

```

```
Query Match      64.5%; Score 2227.2; DB 1; Length 3501;
Best Local Similarity 79.6%; Pred. No. 0;
Matches 2775; Conservative 0; Mismatches 653; Indels 56;
```

Qy	8	GGCTACCGCGCGCTCCGAGGCGCATGCGGGCGTTGGCGCAGCGACGCGCGGCTGC	67
Db	35	GGGACCCCGGCTCCGAGGCGCATGCGGGCGTTGGCGCAGCGCGGCAACCGTGC	91
Qy	68	TGCTGCTTGTGTTTGTCACTAATGATCTTTGAGACCGTTACAAACCAAGACCTGCCTG	127
Db	92	CGCTGCTGCTGTTTCTGCAATGATATTTGGGACTATTACAAATCAAGATCTGCCTG	151
Qy	128	TGATCAAGTGTTTAAATCAGTCATGAGACAATGGCTCATCAGCGGGAAGCATCAT	187
Db	152	TGATCAAGTGTTTAAATCAATCAATAGACAATGATTCATCAGTGGGGAAGTCAATCAT	211
Qy	188	CGTACCGAATGGTGGAGGATCCCGAGAGACCTCCAGTGATACCCGAGGCGCCAGATG	247
Db	212	CATATCCATGTTATCAAAATCCCGGAAGACCTCGGGTGTGCTTGTAGACCCCGAGGCT	271
Qy	248	AAGGGAACGGTATATGAAGCGGCCACCGTGGAGGTGGCGGAGTCTGGGTCCATCAACCTGC	307
Db	272	CAGGACAGTGTACGAAGCTGCCGTGTGGAGTGGATGATCTGCTCTCATCACTATGC	331
Qy	308	AAGTGCAGCTCGCAACCCCGAGGACCTTTCTGCTCTGGGTCTTTAAGCACAGCTCCC	367
Db	332	AAGTGTGGTGGATGCGCCCGAGGAACATTTCTGCTCTGGGTCTTTAAGCACAGCTCCC	391
Qy	368	TGGGCTGCCAGCCGACTTTGATTTACAAAACAGAGGAATGTTTTCATGGCCATCTTGA	427
Db	392	TGAATTGCCAGCCACATTTTGAATTTACAAAACAGAGGAGTTGTTTTCATGGTCAATTTGA	451
Qy	428	ACGTGACAGAGACCCAGCGAGGAAATACCTACTCATTTTCAGAGCGCAACGCGCAACT	487
Db	452	AAATGACAGAAACCCAGCTGGAGAAATACCTACTTTTATTTCAGAGTGAAGTACCAATT	511
Qy	488	ACACAGTACTGTTCCACAGTGAATGTAAAGATACACAGCTGTATGCTTAAGAGACCTT	547
Db	512	ACACAAATATTGTTTACAGTGAATTAAGAAATACCTGCTTTACATTTAAGAGACCTT	571
Qy	548	ACTTTAGGAAGATGAAAACAGGATGCACTGCTCTGCATCTCCGAGGTGTTCCGAGC	607
Db	572	ACTTTAGAAAAATGAAAAACAGGACGCCCTGGTCTGCATATCTGAGAGCGTTCCAGAGC	631
Qy	608	CCACTGTGGATGGGTGCTCTGACGCTCCCAACAGGAAAGCTGTAAAGAAAGCCCTG	667
Db	632	CGATCGTGGAAATGGGTGCTTTGCGAATTCACAGGGGGAAGCTGTAAAGAAAGTCCAG	691
Qy	668	CTGTTGTGAAAGGAGGAAGTACTTTATGATGTTTTCGGAACACACATCATGATGCT	727
Db	692	CTGTTGTAAAAAGGAGGAAGATGCTTCTATGAAATTTATTTGGGACGGAATGAGTGTCT	751
Qy	728	GTGCTAGAAATGCACTGGCGCGCAATGCAACCAAGCTTTCACCATAGATCTTAACACAG	787
Db	752	GTGCCAAGAAATGAATCTGGGACGGAAATGCAACAGGCTGTTTACAAATAGATTAATACAA	811
Qy	788	CTCCTCAGAGCACACTGCCCGCAGTTATTTCTGAAAGTGGGGGAAACCCCTGTGGATCAGT	847

812	Db	CTCCTCAGACCAATTCGCCAATATTATTTCTTAAAGTAGGGAAACCCCTTATGGATAAGGT	871
848	Qy	GTAAGGCCATCCCATGTCAACCATGGATTTCGGGCTCACCTGGAGCTCGAAGACAAAGCCC	907
872	Db	GCAAGCTGTTTCATGTGAACCATGGATTTCGGGCTCACCTGGGAATTAGAAAACAAGCAC	931
908	Qy	TGGAGGAGGCGAGCTACTTTTGAGATGAGTACCTACTCCAAAAACAGACCATGATTCGGA	967
932	Db	TCGAGGAGGCGCACTCTTTGAGATGAGTACCTATTCAACAACAGAACTATGATACGGA	991
968	Qy	TTCTCTTTGGGCTTTTGCTCTTTCGGTGGGAAGGAACGACACGGGATATTACACTGCTCTT	1027
992	Db	TTCTGTTTGCTTTTGTTATCATCATGTGCAAGAAACGACACGGGATCTACACTTGTTCCT	1051
1028	Qy	CCTCAAGACACCCAGCCAGCTCAGCGTTGCTGGACCATCTTAGAAAAAGGGTTTATAAACG	1087
1052	Db	CTTCAAAAGCATCCCGAGTCAATCAGCTTTGTTTACCATCTGTAGGAAGGGAATTTATAAATG	1111
1088	Qy	CTACGAGCTCGCAAGAGAGTATGAATTTGACCCGTACGAAAAGTTCTGCTTCTCAGTCA	1147
1112	Db	CTACCAATTCAGTGAGATTATGAATTTGACCAATATGAAGAGTTTGTGTTTCTCTGTC	1171
1148	Qy	GGTTTAAAGCGTACCCACGAATTCGATGCAAGTGGATCTTCTCTCAAGCCCTCATTTCCCTT	1207
1172	Db	GGTTTAAAGCGTACCCACGAATCAGATGTTACGTGGACCTTCTCTCGAAAAATCATTTCCCTT	1231
1208	Qy	GTGAACAGAGAGCGCTCGAGGATGGGTACAGCATATCTAAATTTTGGCATCATGAAGAACA	1267
1232	Db	GTGAGCAAAAGGGCTTTGATAACGGATACAGCATATCCAAGTTTGGCAATCATAAAGCAC	1291
1268	Qy	AGCCAGGAGAGTACATATTTCTATGCAGAAATGATGACGCCAGTTTCAACCAAAATGTTCA	1327
1292	Db	AGCCAGGAGAAATATATTTCAATGCAGAAATGATGATGCCAATTTACCAAAATGTTCA	1351
1328	Qy	CGCTGAATATAAGAAAGAAACCTCAAGTGTAGCAAAATGGCTCAGCCAGCCAGCGCTCCT	1387
1352	Db	CGCTGAATATAAGAAAGAAACCTCAAGTGTCTGCAGAAAGCATCGGCAAGTCAGGCGTCT	1411
1388	Qy	GTTCCTCTGATGCTACCGCTACCTCTTGGACCTTGGAGAGAGTGTTCGGACAAATCTC	1447
1412	Db	GTTTCTCGGATGATACCCATTACCATCTTTGACCTTGGAAAGAGTGTTCAGACAAGTCTC	1471
1448	Qy	CCAATTGACGAGGAAATCCCAAGAGGAGTTTGGAAATAAAAGGCTAAACAGAAAAAGTGT	1507
1472	Db	CCAATGCAAGAGAGATCAGAGAGGAGTCTGGAAATAGAAAGGCTTAACAGAAAAGTGT	1531
1508	Qy	TTGGCCAGTGGGTGTGAGCAGTACTCTAAATATGATGAGCGCGGAAAGGGCTTCTGG	1567
1532	Db	TTGCAAGTGGGTGTGAGCAGTACTCTAAACATGATGTAAGCCATCAAAAAGGGTCTCTGG	1591
1568	Qy	TCAAAATGCTGCGTACAATTTCTATGGCACGCTTTGGCAACCATCTTTTAAACCTCAC	1627
1592	Db	TCAAGTGCTGTGCATACAAATTCCTTTGGCACATCTTTGTGAGACGATCTTTTAAACCTC	1651
1628	Qy	CAGGCCCCCTTCCCTTTTCATCCAAGACAAACATCTCTTCTATGCGACCATTTGGGCTCTGTC	1687
1652	Db	CAGGCCCCCTTCCCTTTTCATCCAAGACAAACATCTCATTTCTATGCAACAAATTTGGTTTGT	1711
1688	Qy	TCCCTTTCATTTGTTTCTCATTTGTTGATCTGCAACAAATACAAAAGCAATTTAGGT	1747
1712	Db	TCCTCTTCTTGTTCGTTTAAACCTCTCTAAATTTGTCAAGTACAAAAGCAATTTAGGT	1771
1748	Qy	ACGAGAGTCAGCTGCAGATGATCCAGGTGATCGGGCCCCCTGGATATACGATGATCTTACG	1807
1772	Db	ATGAAAGCCAGCTACAGATGGTACAGGTGACCGGGCTCTCAGATATATGATGATCTTACG	1831
1808	Qy	TTGACCTTCAGGAGTATGAATATGACCTTAAAGTGGGAGTTCCGAGAGAGAACTTAGGT	1867
1832	Db	TTGATTTCAAGANATATGATATGATCTCAATGGGAGTTTCAAGAGAAAATTTAGAGT	1891
1868	Qy	TTGGGAAGGTCTCGGGCTTGGCGCTTTCGGAGGGTGATGAACGCGACGCGCTATGGCA	1927

1892 TTGGGAAGGTACTAGGATCAGGTGCTTTTGGAAAAGTGATGAACGCAACAGCTTATGGAA 1951
1928 TTAGTAAACGGGAGCTCTCAATTCAGGTGGCGTGAAGATGCTAAAGAGAAAGCTGACA 1987
1952 TTAGCAAAACAGGAGTCTCAATTCAGGTGGCGTCAAAATGCTGAAGAAAAGACAGACA 2011
1988 GCTGTGAAAAGAAAGCTCTCATGTGGAGCTCAAAATGATGACCCACCTGGGACACCATG 2047
2012 GCTCTGAAAGAGAGCACTCATGTGAGAACTCAAGATGATGACCCAGCTGGGAGCCACG 2071
2048 ACAACATCGTAATCTGTGGGGGATGACACTGTGAGGCCAGTGTACTTGAATTTTGG 2107
2072 AGAATAATGTGAACCTGTGTGGGGGCTGACACTGTGAGGACCAATTTACTTGAATTTG 2131
2108 AATATTGTTGTATGTGACCTCTCAACTACCTTAAGAACTCAAGATGATGAGAGAGTTCACA 2167
2132 AATACTGTTGTATGTGATCTTCAACTATCTAAGAAAGTAAAGAGAGAAAATTTTACA 2191
2168 GGACATGGACAGAGATTTTAAAGAAACATAATTTTCACTTCTTACCTTCTTCCAGGCAC 2227
2192 GGACTTGGACAGAGATTTTCAAGGAACACAAATTTTCACTTCTTACCTTCTTCCATCAC 2251
2228 ATTCAATTTCCAGATGCTGTGGTTCAGAGAAGTTCAGATTACACCCGCTTGGATCAGC 2287
2252 ATCCAAATTCAGCATGCTGTGTTCAAGAGAAGTTCAGATACACCCGCACTCGGATCAA 2311
2288 TCTCAGGGTTCATGGGAATTCATTTCTGTAAGATGAGATGAAATGAAGAAACGAGA 2347
2312 TCTCAGGGCTTCATGGGAATTCATTTCACTCTGAAGATGAAATGAAATGAAGAAACCAA 2371
2348 AGAGGCTGGCAGAGAAGAGAGAGAGATTTGAACGTGCTGACGTTTGAAGACCTCTCTT 2407
2372 AAAGGCTG-----GAAGAGAGAGAGACTTGAATGTGCTTACATTTGAAGATCTCTTT 2425
2408 GCTTTGCGTACCAAGTGGCCAAAGGCAATGGAATTCCTGGAGTTCAAGTCTGTGTCCACA 2467
2426 GCTTTGCATATCAAGTGTGCCAAAGGAATGGAATTTCTGGAAATTAAGTCTGTGTTCACA 2485
2468 GAGACCTGGCAGCCAGGAATGTGTGTGTCACCCAGGGAAGGTGGTGAAGATCTGTGACT 2527
2486 GAGACCTGGCGCCGAGGAACGTGTTGTCACCCAGGGAAGGTGGTGAAGATATGTGACT 2545
2528 TTGGACTGGCCCGAGACATCTCAGCGACTCCAGCTACGTCGTCAGGGGCAACGCAAGGC 2587
2546 TTGGATTGCTCGAGATATCATGAGTGAATTCGAACATATGTTGTCAGGGGCAATGCCCGTC 2605
2588 TGCCGCTGAAGTGAATGGCAACCGAGAGCTTATTTGAAGGGAATCTACACAATCAAGAGTG 2647
2606 TGCCCTGTAATAATGGATGGGCCCGCAAGGCTGTTTGAAGGCACTACACCATTAAGAGTG 2665
2648 ACGTCTGGTCTTACGCACTCTTCTCTGGAGATATTTTCACTGGGTGTGAACCCCTTACC 2707
2666 ATGCTGGTCAATAGGAATATTAATCTGGGAAATCTTCACTTGGTGTGAATCTCTTACC 2725
2708 CTGGCAATCTCTGTGACGCTAACTCTATAAATGATTCAGAGTGAATTTAAATGGAGC 2767
2726 CTGGCAATCTCTGTGATGTAATCTTCAAACTGATTCAAAATGGAATTTAAATGGATC 2785
2768 AGCCATTTATGCCAGAGAGGATATATCTTTGTAATGCAATCTCTGCGGCTTTTGGACT 2827
2786 AGCCATTTATGCTACAGAGAATAATACATTAATGCAATCTCTGCGGCTTTTGGACT 2845
2828 CAAGGAAGGGCCATCTTCCCAACCTGACTTCAATTTTGAATGTGCTGCGGAGAGG 2887
2846 CAAGGAAGGGCCATCTTCCCAATTTGACTTCGTTTGAAGTGTGCTGCGGAGAGG 2905
2888 CAGAAGAAG-----CATGTATCAAAATCCATCCATCTACC 2924
2906 CAGAAGAAGCGATGTATCAGATGTGGATGGCGGTGTTTGGGAATGTCTCTCACCTTACC 2965
2925 AAAACAGCGGCCCTCTCAGCAGAGAGGGGGCTCAGAGC-CCAGTCCGACACAGCGCCAGG 2983
2966 AAAACAGCGGCCCTTTCAGCAGAGAGATGGATTTGGGGCTACTCTCTCGCAGGCTCAGG 3025

QY 2984 T-GAAGATTCA CAGAGAAAGATTAGCGAGGAGCCTTTGACCCCGCC-----ACCTAG 3038
D 3026 TCGAAGATTCTG TAGAGGAACAATTTAGTTTAAAGGACTTCATCCCTCCACCTATCCCTAA 3085
QY 3039 CAGGCTGTAGACCGCAGAGCCAGAGATTAGCCTCGCCTCT--GAGGAAGCGCCCTACAGCG 3096
D 3086 CAGGCTGTAGATTACCAAAAACAAGATTAAATTTTCATCACTAAAAGAAAATCTATTATCAAC 3145
QY 3097 CGTTGCTTCGCTGGACTTTTCTCTAGATGCTGTCTGCCATTACTC-----CAAAGTGA 3149
D 3146 TGCTGCTTCCACAGACTTTTCTCTAGAGCCGCTCTGCGTTTACTCTTGTGTTTCAAGGGA 3205
QY 3150 CTTCTATAAATCAAACTCTCTCGCACAGCGCGGAGAGCAATAATGAGACTTGTGG 3209
D 3206 CTTTGTGTAATAATCAAAATCATCTCTGCACAGGCGAGGAGACTGATAATGAACTTTATG 3265
QY 3210 TGACCCCGCTACCTCGGGGCTTTTCCACGAGCTTGAGGGGAAGCCATGTATCTGAAA 3269
D 3266 GAGCAATTGATCTGCAATCCAAAGGCTTCTCAGGCGCGCTTGAGTGAATGTGTACCTGAAG 3325
QY 3270 TATAGTATATTTCTGTAAATACGTTGAAACAAACCAACCCGCTTTTTCCTAAGGGAAGC 3329
D 3326 TACAGTATATTTCTGTAAATACATAAAACAAA-----AGCATTTTGTCTAAGGAGAAC 3378
QY 3330 TAAATATGATTTTAAATACTATGTTTAAATACTATGTAACTTTTTCATCTATTATTAG 3389
D 3379 TAATATGATTTT--AAGTCTATGTTTAAATAATATATGTAATTTTTCAGCTATTATTAG 3436
QY 3390 TGATATATTTATGGATGGAATAAATCTTCTACTGTAAATAAATAAATAAATAAATAA 3449
D 3437 TGATATATTTTATGGGTGGGAATAAATAATTTCTACTACAGAAAATAAATAAATAA 3496
QY 3450 AAAA 3453
D 3497 AAAA 3500

Search completed: August 28, 2003, 05:51:50
Job time : 296.55 sec

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2003, 07:20:58 ; Search time 32.9834 Seconds
(without alignments)
2892.346 Million cell updates/sec

Title: US-09-919-408A-2

Perfect score: 5264

Sequence: 1 MRALAQRSDRLLLVLSV.....RGLRAQSPQRQVQKIHRS 992

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5264	100.0	992	2	protein-tyrosine k
2	5102	96.9	1000	2	Flt3 protein - mou
3	4408.5	83.7	993	2	protein-tyrosine k
4	1258.5	23.9	980	1	macrophage colony
5	1247	23.7	941	1	protein-tyrosine k
6	1232.5	23.4	978	1	protein-tyrosine k
7	1223	23.2	975	1	protein-tyrosine k
8	1221.5	23.2	972	1	macrophage colony
9	1220	23.2	954	2	c-kit-related kina
10	1215	23.1	977	2	protein-tyrosine k
11	1210	23.0	978	2	macrophage colony
12	1198.5	22.8	976	1	macrophage colony
13	1196	22.7	976	1	protein-tyrosine k
14	1188.5	22.6	975	2	macrophage colony
15	1181	22.4	960	1	protein-tyrosine k
16	1157.5	22.0	1088	1	platelet-derived g
17	1152.5	21.9	1089	1	platelet-derived g
18	1132	21.5	1087	2	platelet-derived g
19	1123.5	21.3	1089	1	platelet-derived g
20	1098	20.9	1098	1	platelet-derived g
21	1078	20.5	1106	1	platelet-derived g
22	1060	20.1	1048	2	platelet-derived g
23	998.5	19.0	1338	2	protein-tyrosine k
24	984.5	18.7	1336	2	Fit-1 tyrosine kin
25	983.5	18.7	1333	2	receptor tyrosine
26	980	18.6	1379	2	vascular endotheli
27	970	18.4	1330	2	embryonic receptor
28	967.5	18.4	790	1	gag-kit polyprotei
29	965	18.3	1348	2	vascular endotheli

30 954.5 18.1 1298 2 A48999 protein-tyrosine k
31 954.5 18.1 1356 2 JC1402 protein-tyrosine k
32 952 18.1 1363 2 I58375 protein-tyrosine k
33 926.5 17.6 1367 2 A41228 protein-tyrosine k
34 822 15.6 160 2 A39061 fibroblast growth
35 815 15.5 813 1 A49123 protein-tyrosine k
36 810.5 15.4 823 2 B35963 protein-tyrosine k
37 798 15.2 797 2 S38579 fibroblast growth
38 791 15.0 821 1 TVHUF2 fibroblast growth
39 791 15.0 824 2 S24108 protein-tyrosine k
40 789 15.0 806 2 A35963 protein-tyrosine k
41 788.5 15.0 769 2 S16236 fibroblast growth
42 788.5 15.0 822 2 A45081 fibroblast growth
43 780.5 14.8 822 2 A41794 keratinocyte growt
44 779.5 14.8 457 2 S44269 platelet-derived g
45 777 14.8 821 1 TVMSBK fibroblast growth

ALIGNMENTS

RESULT 1

A39931
protein-tyrosine kinase (EC 2.7.1.112) flk-2, hematopoietic - mouse
C/Species: Mus musculus (house mouse)
C/Date: 17-Jan-1992 #sequence_revision 17-Jan-1992 #text_change 04-Feb-2000
C/Accession: A39931
R/Matthews, W.; Jordan, C.T.; Wiegand, G.W.; Pardoll, D.; Lemischka, I.R.
Cell 65, 1143-1152, 1991
A/Title: A receptor tyrosine kinase specific to hematopoietic stem and progenitor cell-
A/Reference number: A39931; MUID:91292518; PMID:1648448
A/Accession: A39931
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-992 <NAT>
A/Cross-references: GB:M64689; NID:q193327; PIDN:AAA37634.1; PID:q193328
C/Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homol-
C/Keywords: ATP; phosphotransferase; transmembrane protein; tyrosine-specific protein k
F/609-953/Domain: protein kinase homology <KIN>
F/617-625/Region: protein kinase ATP-binding motif

Query Match 100.0%; Score 5264; DB 2; Length 992;

Best Local Similarity 100.0%; Pred. No. 8.6e-263; Mismatches 0; Indels 0; Gaps 0;
Matches 992; Conservative 0;

Qy	1	MRALAQRSDRLLLVLSV	WMILETVNQDLPVTKVLI	SHENNGSSAGKPS	SYRMVRGS	60																																												
Db	1	MRALAQRSDRLLLVLSV	WMILETVNQDLPVTKVLI	SHENNGSSAGKPS	SYRMVRGS	60																																												
Qy	61	PEDLQCTPRQSEGT	VYEAATVEA	SGSITLQV	LATPGDLSCLW	FKHSSLCGQPHFD	120																																											
Db	61	PEDLQCTPRQSEGT	VYEAATVEA	SGSITLQV	LATPGDLSCLW	FKHSSLCGQPHFD	120																																											
Qy	121	LQNRGVSMALNV	TQAGEYLLH	IQSERANY	TVLFTVNV	RDTQVLT	VPYRKME	180																																										
Db	121	LQNRGVSMALNV	TQAGEYLLH	IQSERANY	TVLFTVNV	RDTQVLT	VPYRKME	180																																										
Qy	181	DALLCISEGV	PEPTVEW	LVLCSSH	RESCKEG	PAVRKEE	KVLH	ELFGTDRCCARNALGR	240																																									
Db	181	DALLCISEGV	PEPTVEW	LVLCSSH	RESCKEG	PAVRKEE	KVLH	ELFGTDRCCARNALGR	240																																									
Qy	241	ECTKLF	TIDL	NAQPS	TLP	FLKVG	EPLW	IRCKAI	HVNHG	FGLTWE	DKAL	EGSYFE	300																																					
Db	241	ECTKLF	TIDL	NAQPS	TLP	FLKVG	EPLW	IRCKAI	HVNHG	FGLTWE	DKAL	EGSYFE	300																																					
Qy	301	MSY	STN	TR	TMIR	ILL	AP	VSS	GR	ND	TY	TCSS	KHP	SQ	SAL	VT	ILE	K	G	F	N	A	T	S	S	O	B	E	E	Y	420																			
Db	301	MSY	STN	TR	TMIR	ILL	AP	VSS	GR	ND	TY	TCSS	KHP	SQ	SAL	VT	ILE	K	G	F	N	A	T	S	S	O	B	E	E	Y	420																			
Qy	361	EIDPY	E	K	F	C	F	S	V	R	K	A	P	R	I	C	T	W	I	S	F	S	A	P	C	E	O	R	G	L	E	D	G	Y	S	I	S	K	F	C	D	H	N	K	P	E	G	E	Y	420
Db	361	EIDPY	E	K	F	C	F	S	V	R	K	A	P	R	I	C	T	W	I	S	F	S	A	P	C	E	O	R	G	L	E	D	G	Y	S	I	S	K	F	C	D	H	N	K	P	E	G	E	Y	420

Qy	421	AENDDAQTKMFTLNIRKKPQVLANASQAQSSDGYPLPSWTWKKCDKSPNCTTEIIP	480
Db	421	AENDDAQTKMFTLNIRKKPQVLANASQAQSSDGYPLPSWTWKKCDKSPNCTTEIIP	480
Qy	481	EGVWNKKANRVKFGQWVSSSTLNMSEAGKGLLVKCCAYNSMGTSCETIFLNSPGPPFPQ	540
Db	481	EGVWNKKANRVKFGQWVSSSTLNMSEAGKGLLVKCCAYNSMGTSCETIFLNSPGPPFPQ	540
Qy	541	DNISFYATIGLCLPFIIVLVILVLI	600
Db	541	DNISFYATIGLCLPFIIVLVILVLI	600
Qy	601	DLKWEFPRENLEFGKVLGSGAFGRVNNATAYAGISKTGVSIVAVXMKLKEKADSCKEALM	660
Db	601	DLKWEFPRENLEFGKVLGSGAFGRVNNATAYAGISKTGVSIVAVXMKLKEKADSCKEALM	660
Qy	661	SELKXMTHLGHHNDINVLGACTLSGPPVLI	720
Db	661	SELKXMTHLGHHNDINVLGACTLSGPPVLI	720
Qy	721	EHNFSYPTFOAHNSNMPGSRVQLHPDQLSGFNGNSIHSEDEIEYENOKRLAESEE	780
Db	721	EHNFSYPTFOAHNSNMPGSRVQLHPDQLSGFNGNSIHSEDEIEYENOKRLAESEE	780
Qy	781	EDLNVLTPEDELLCFAYQVAKGMFEFLFKSCVHRDLAARNVLVTHGKVVKICDFGLARDIL	840
Db	781	EDLNVLTPEDELLCFAYQVAKGMFEFLFKSCVHRDLAARNVLVTHGKVVKICDFGLARDIL	840
Qy	841	SDSSVYVVRGNARLPVKWMAPESLFPGIYTIKSDVWSYGILLWEIFSLGWNVPYGPDPVAN	900
Db	841	SDSSVYVVRGNARLPVKWMAPESLFPGIYTIKSDVWSYGILLWEIFSLGWNVPYGPDPVAN	900
Qy	901	FYKLIQSGFKMEQFPYATYEGYFVVMQSWAFDSRKRPSFPNLTSLFLCOLAABEACIRT	960
Db	901	FYKLIQSGFKMEQFPYATYEGYFVVMQSWAFDSRKRPSFPNLTSLFLCOLAABEACIRT	960
Qy	961	SIHLPKQAAPQOQGLRAQSPQKVIHRS	992
Db	961	SIHLPKQAAPQOQGLRAQSPQKVIHRS	992
RESULT 2			
S18827			
Flt3 protein - mouse			
C;Species: Mus musculus (house mouse)			
C;Date: 19-Feb-1994 #sequence_revision 10-Nov-1995 #text_change 24-Sep-1999			
C;Accession: S18827			
R;Rosnet, O.; Marchetto, S.; delapeyriere, O.; Birnbaum, D.			
Oncogene 6, 1641-1650, 1991			
A;Title: Murine Flt3, a gene encoding a novel tyrosine kinase receptor of the			
A;Reference number: S18827; PMID:1656368			
A;Accession: S18827			
A;Molecule type: mRNA			
A;Residues: 1-1000 <ROS>			
A;Cross-references: EMBL:X59398; NID:g50978; PIDN:CAA42041.1; PID:g50979			
C;Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kin			
C;Keywords: ATP			
F:609-953/Domain: protein kinase homology <KIN>			
F:617-625/Region: protein kinase ATP-binding motif			
Query Match 96.9%; Score 5102; DB 2; Length 1000;			
Best Local Similarity 96.7%; Pred. No. 1.8e-254;			
Matches 968; Conservative 6; Mismatches 17; Indels 10; Gaps 2			
Qy	1	MRALAQRSDRLLLLVLSVMILETVTNQDLPVVKCVLISHNNNGSSAGKPSYRWGRS	60
Db	1	MRALAQRSDRLLLLVLSVMILETVTNQDLPVVKCVLISHNNNGSSAGKPSYRWGRS	60
Qy	61	PEDLQCTPRROSEGTIVYEAATVEVAESGITLQVLATPGDLSCLVFKHSLGCGQPHD	120
Db	61	PEDLQCTPRROSEGTIVYEAATVEVAESGITLQVLATPGDLSCLVFKHSLGCGQPHD	120
Qy	121	LQNRGIVSMALNVTQAGBYLLHIQSERANYTLVFTVNVKDTQLVYLRFRFFKMKQ	180

Db	121	LQNRGISMAI	N	T	T	E	T	O	A	G	S	Y	L	L	H	I	O	S	E	A	N	Y	T	V	L	F	N	V	R	D	T	O	L	Y	L	R	R	P	F	R	K	M	E	N	Q	180						
Qy	181	DALLCISG	V	P	E	P	T	V	E	W	L	C	S	S	H	R	E	S	C	K	E	E	G	P	A	V	R	K	E	K	V	L	H	E	L	F	G	T	D	I	R	C	C	A	R	N	A	L	G	R	240	
Db	181	DALLCISG	V	P	E	P	T	V	E	W	L	C	S	S	H	R	E	S	C	K	E	E	G	P	A	V	R	K	E	K	V	L	H	E	L	F	G	T	D	I	R	C	C	A	R	N	A	L	G	R	240	
Qy	241	ECTKLFTI	D	L	N	O	A	P	O	S	T	L	P	Q	L	F	K	V	G	B	E	L	R	I	C	K	A	I	H	V	N	H	G	F	L	T	W	E	D	K	A	L	E	G	S	Y	F	E	300			
Db	241	ESTKLFTI	D	L	N	O	A	P	O	S	T	L	P	Q	L	F	K	V	G	B	E	L	R	I	C	K	A	I	H	V	N	H	G	F	L	T	W	E	D	K	A	L	E	G	S	Y	F	E	300			
Qy	301	MSTYSTNR	T	M	T	R	I	I	L	A	F	V	S	S	V	G	R	N	D	T	G	Y	T	C	S	S	K	H	P	S	Q	S	A	L	V	T	L	E	K	G	F	I	N	A	T	S	S	O	B	E	Y	360
Db	301	MSTYSTNR	T	M	T	R	I	I	L	A	F	V	S	S	V	G	R	N	D	T	G	Y	T	C	S	S	K	H	P	S	Q	S	A	L	V	T	L	E	K	G	F	I	N	A	T	S	S	O	B	E	Y	360
Qy	361	EIDPYE	K	F	C	F	S	R	V	R	K	A	P	R	I	R	C	T	M	I	F	S	O	A	S	P	C	E	O	R	G	L	E	D	G	Y	S	I	S	K	F	C	D	H	K	N	K	P	E	Y	420	
Db	361	EIDPYE	K	F	C	F	S	R	V	R	K	A	P	R	I	R	C	T	M	I	F	S	O	A	S	P	C	E	O	R	G	L	E	D	G	Y	S	I	S	K	F	C	D	H	K	N	K	P	E	Y	420	
Qy	421	AENDDAQ	T	K	M	F	T	L	N	I	R	K	P	O	V	L	A	N	A	S	A	S	O	A	S	C	S	D	G	Y	L	P	S	W	T	W	K	K	C	S	D	K	S	P	N	C	T	E	E	P	480	
Db	421	AENDDAQ	T	K	M	F	T	L	N	I	R	K	P	O	V	L	A	N	A	S	A	S	O	A	S	C	S	D	G	Y	L	P	S	W	T	W	K	K	C	S	D	K	S	P	N	C	T	E	E	P	480	
Qy	481	EGVWNK	A	N	R	V	F	G	O	W	S	S	T	L	N	M	S	E	A	G	K	L	L	V	K	C	A	Y	N	S	M	G	T	S	C	E	T	I	F	L	N	S	P	G	P	P	P	F	I	Q	540	
Db	481	EGVWNK	A	N	R	V	F	G	O	W	S	S	T	L	N	M	S	E	A	G	K	L	L	V	K	C	A	Y	N	S	M	G	T	S	C	E	T	I	F	L	N	S	P	G	P	P	P	F	I	Q	540	
Qy	541	DNISFY	A	T	I	G	L	C	P	F	I	V	L	I	V	L	I	C	H	K	Y	K	Q	R	F	R	Y	E	S	O	L	O	M	I	Q	V	T	G	P	L	D	N	E	Y	F	V	D	R	D	Y	E	600
Db	541	DNISFY	A	T	I	G	L	C	P	F	I	V	L	I	V	L	I	C	H	K	Y	K	Q	R	F	R	Y	E	S	O	L	O	M	I	Q	V	T															

[illegible]

A:Molecule type: mRNA
A:Residues: 1-993 <SMA>
A:Cross-references: GB:002687
A>Note: in the authors translation, an additional residue Ala is shown after 420-Ala and C;Genetics:
A:Map position: 13q12
A:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog
C:Keywords: ATP; glycoprotein; phosphotransferase; transmembrane protein; tyrosine-speci
F:608-950/Domain: protein kinase homology <KIN>
F:616-624/Region: protein kinase ATP-binding motif

Query Match 83.7%; Score 4408.5; DB 2; Length 993;
Best Local Similarity 84.0%; Pred. No. 6.8e-219;
Matches 836; Conservative 57; Mismatches 89; Indels 13; Gaps 6;

QY 1 MRALQSRDRLLLLVLSVILETVNQDLVPIKCVLISHENNGSAGPSPSYRMVGRS 60
DB 1 MPALA-RDAGTVPLLVFSAMIFGTITNQDLVPIKCVLISHENNGSAGPSPSYRMVGRS 59

QY 61 PEDLOTPRQSGTVYEATVEVAESGSLTQVLAATPGDLSCLVFKHSLGCGQPHD 120
DB 60 PEDLGALRPQSGTVYEAAVEVDVASITLQVLDAPGNISCLVFKHSLGCGQPHD 119

QY 121 LQNRGIVSMALNVETQAGEYLLHIOSEANVTYLTQVLRVDTQVLRVPRKVMQ 180
DB 120 LQNRGVVSMVLKMTQAGEYLLHIOSEATNTYLTQVLRVDTQVLRVPRKVMQ 179

QY 181 DALLCISEGVPETVEVWLCSSHRESCKEKPAPVVRKEEVLHFGTDIRCCARNALGR 240
DB 180 DALVCISEGVPETVEVWLCSSHRESCKEKPAPVVRKEEVLHFGTDIRCCARNALGR 239

QY 241 ECTKLTIDLNQAPQSTLPFLKVGPELMIRCKAIVHNVHFGTLTWELEKALBEGYFE 300
DB 240 ECTRLTIDLNQAPQSTLPFLKVGPELMIRCKAIVHNVHFGTLTWELEKALBEGYFE 299

QY 301 MSTYSTNRTMIRLLAFVSSVGRNDTGYTCSSSKHPSQSALVTILEKGFNATSSOEY 360
DB 300 MSTYSTNRTMIRLLAFVSSVGRNDTGYTCSSSKHPSQSALVTILEKGFNATSSOEY 359

QY 361 EIDPYEFCFVPRKAYPRIRCTWIFQAGPPCQRGLEDQYSIKFCDHKNKGEYIFY 420
DB 360 EIDPYEFCFVPRKAYPRIRCTWIFQAGPPCQRGLEDQYSIKFCDHKNKGEYIFY 419

QY 421 -AENDDAQFTKMTLNIRKQVLANASASQSCSDGYPPLPWTWKCDKSPNCTEEI 479
DB 420 AENDDAQFTKMTLNIRKQVLANASASQSCSDGYPPLPWTWKCDKSPNCTEEI 478

QY 480 PEGWVKANRKPQGVQWSSSTLNMSEAGKLLVCCAYNSMGTSCTIFLNSPGPPFI 539
DB 479 TEGWVKANRKPQGVQWSSSTLNMSEAGKLLVCCAYNSMGTSCTIFLNSPGPPFI 538

QY 540 QDNISFATIGLCIPFIVLVILCHYKKQFRYESQLQVQVGTQVLDNEFYVDPDYE 599
DB 539 QDNISFATIGLCIPFIVLVILCHYKKQFRYESQLQVQVGTQVLDNEFYVDPDYE 598

QY 600 YDLKWEPPRENLEFGKVLGSAFGFVNNATAYGISTKGVSIQVAVKMLKEKADSCKEAL 659
DB 599 YDLKWEPPRENLEFGKVLGSAFGFVNNATAYGISTKGVSIQVAVKMLKEKADSCKEAL 658

QY 660 MSELKMTHTLGHNDINVLGACTLSPGVILIFCYCGDGLNLLYLRKREKPHRTWTIEF 719
DB 659 MSELKMTHTLGHNDINVLGACTLSPGVILIFCYCGDGLNLLYLRKREKPHRTWTIEF 718

QY 720 KEHNFSSYPPTFAHNSMPSGSRVQLHPPDQLSGFNGNSIHSEDELEYENQRLAEE 779
DB 719 KEHNFSSYPPTFAHNSMPSGSRVQLHPPDQLSGFNGNSIHSEDELEYENQRLAEE 776

QY 780 EEDNLVLTFFEDLLCFAYQVAKGMEFLBFKSCVHRDLAARNVLTGHKVKICDFGLARDI 839
DB 777 EEDNLVLTFFEDLLCFAYQVAKGMEFLBFKSCVHRDLAARNVLTGHKVKICDFGLARDI 836

QY 840 LQDSSVYVRGNARLPVKWMAPESEFEGIYTIKSDVWSYGILLWEIFSLGVNYPGIPVDA 899
DB 837 MSDSNYVVRGNARLPVKWMAPESEFEGIYTIKSDVWSYGILLWEIFSLGVNYPGIPVDA 896

QY 900 NFYKLTQSGFKMQPFPVATGIIYVWQSCWAFOSRKPSPNLTSLFGCCQLABEAC-- 957
DB 897 NFYKLTQSGFKMQPFPVATGIIYVWQSCWAFOSRKPSPNLTSLFGCCQLABEANYQ 956

QY 958 -----TSTSHLPKQAAPOQRG-GLRAQSPQKQV 986
DB 957 NVDRVSECHTYQNRPPSPREMDLGLLSPOAQVE 991

RESULT 4

TVCTMD

macrophage colony-stimulating factor 1 receptor precursor - cat

N:Contains: protein-tyrosine kinase (EC 2.7.1.12) csfir/fms

C:Species: Felis silvestris catus (domestic cat)

C>Date: 31-Dec-1989 #sequence_revision 31-Dec-1989 #text_change 13-Jun-1997

C:Accession: A31636

R:Woolford, J.; McAuliffe, A.; Rohrschneider, L.R.

Cell 55, 965-977, 1988

A:Title: Activation of the feline c-fms proto-oncogene: multiple alterations are requir

A:Reference number: A31636; MUID:89077553; PMID:2849512

A:Accession: A31636

A:Molecule type: mRNA

A:Residues: 1-980 <WOO>

A:Cross-references: EMBL:X03663

C:Genetics:

A:Gene: fms

C:Superfamily: macrophage colony-stimulating factor 1 receptor; immunoglobulin homology

C:Keywords: Anp; autophosphorylation; glycoprotein; kinase-related transforming protein

f1c protein kinase

F:1-23/Domain: signal sequence #status predicted <SIG>

F:24-980/Product: macrophage colony-stimulating factor 1 receptor #status predicted <WA

F:24-509/Domain: extracellular #status predicted <EXT>

F:35-86/Domain: immunoglobulin homology <IMM1>

F:120-179/Domain: immunoglobulin homology <IMM2>

F:217-280/Domain: immunoglobulin homology <IMM3>

F:316-381/Domain: immunoglobulin homology <IMM4>

F:410-484/Domain: immunoglobulin homology <IMM5>

F:535-980/Domain: transmembrane #status predicted <TM>

F:577-915/Domain: intracellular #status predicted <INT>

F:585-593/Region: protein kinase ATP-binding motif

F:42-84,127-177,224-278,417-482/Disulfide bonds: #status predicted

F:45,73,94,153,275,302,335,410,477,490/Binding site: carbohydrate (Asn) (covalent) #sta

F:613,630,776/Active site: Lys, Glu, Asp #status predicted

F:781,794/Binding site: magnesium (Asn, Asp) #status predicted

Query Match 23.9%; Score 1258.5; DB 1; Length 980;

Best Local Similarity 32.7%; Pred. No. 2.9e-57;

Matches 328; Conservative 168; Mismatches 353; Indels 153; Gaps 32;

QY 75 TVVEAATVEVAESGSLTQVLAATPGDLSCL-----W---VFKHSSIGCOPHPDLQNRG 125

DB 13 TAAHAQGVPIQSGPELVPEPTVTLRCVNGSVWDGPISPHNLDDPPSSI-----68

QY 126 IVSMAILNVETQAGEYLLHI-----QSERANYTLFTVNVNRTQVLYL--RRPY-----173

DB 69 ---LTTNNAIFONTGTY--HCTEPNPGGNATI-----HLVYKDPARPKVLAQE 114

QY 174 FRKMNODALL-CISEGVPETVEVWLCSSHRESCKEKPAPVVRKEEVLHFGTDI--230

DB 115 VTVLEGQDALLPCL---LTPDALE-----AGSVLRVRGRPLVQTQYSPFPGHFTTHK 166

QY 231 -----RCCANALGRECTKL-----FTIDLNAQPSLT--POLFLK--GEPLWIRCK 274

DB 167 AKFTENHYQCSAR-VDRGTVTSMGIWLVKQKQISGPAATLLEPAELVRIQEAQAIVCS 225

QY 275 AIHNVHFGTLTWELEKALBEGSYFEMSTYSTNRTMIRLLAFVSSVGRNDTGYTCSSS 334

DB 226 ASNIDNVFVSLRHGDTKL---TISQSDFDHNYQ-KVLTNLNLDHVSFDAGNYSCTAT 281

QY 335 K---HPSQSALVTILEKGFNATSSQE-EYEIDPYEFCFVPRKAYPRIRCTWII-----385

```

Db 282 NAWGNHSAFVRVVEAYLNLTSQSLLEQVTVGEKVDLQVKKVAYPGLESFNNYTLGP 341
Qy 386 FSQASFPCEORLENGYSKPCD-----HKNKPGYIFVAENDDAQFTQVFNIRKPKQ 441
Db 342 FSDYQDKLDFVTKYTRYSLSLPRKRSAGYSFLARNAGQONALTPELLTLYRYPE 401
Qy 442 VLANASQAS-----CSSDGYPLPSWTKKCDKSPNCTEE---IPEGVNNKANKRVFG 494
Db 402 VRVTMLINGSDTLLCEASGYQPSVTVVQCRSHTRDCDEAGLVLESHSEVLQSVPFH 461
Qy 495 QWSSSTLNMSAGKGLLVKCCAYNMGTSCTETFLNSPGPPFFIODNISFYATIGLC-- 552
Db 462 EVIVHSLAIGTLEHNRVTECFRNFVSGNSQTFWPISIGAHTQPLDELLFPVLLTCMS 521
Qy 553 -LPFTIVLVILCHYKQKQRYESQLOMIVTGLDNEYFYVDFRDYDLKWEPPRENIL 611
Db 522 IMALLLLLLLLLYKQKPKYQVRWKIE--SYEGNSYTFIDPTQLPNEKWEPPRNIL 579
Qy 612 EFGKVLGSAGFRVNNATAYGISTGVSIOVAVKMLKEKADSCKEALMSELKMMTHLGH 671
Db 580 QFGKTLGAGAFGVVEATAFGLGKEDAVLKAVKMLKSTAHADKEALMSELKIMSHLQ 639
Qy 672 HDNIVNLGACTLSPGVVLIPEYCCYGDLLNLYRSKRE-----KF 711
Db 640 HENIVNLGACTHGGPVLIVTEYCCYGDLLNLYRSKRE-----KF 711
Qy 712 HRTWTEIFKEHNFSSYPTFOAHNSMPCSRVQLHPPLDQSGFNGSIHSEDEIYEN 771
Db 700 HLEKYYVRDSDFS-----SQVDYTVEMRPVSTSSNDSEEDL----- 740
Qy 772 QKRLAEBEEDNLVITFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLTGHKVKIC 831
Db 741 -----GKEDGRPLELRLDLHFSQVAGQMAFLASKNCIHRDVAARNVLTSGRVAKIG 793
Qy 832 DFLGLARDILSDSSVYVRGNARLPVKWAPESLFEIGIYIKSDVWSYGILLWEIFSLGVP 891
Db 794 DFLGLARDILSDSSVYVRGNARLPVKWAPESLFEIGIYIKSDVWSYGILLWEIFSLGVP 853
Qy 892 YPGIPVDANFYKLIQSGFKMEOPFVATGIVFMQSCWAFSRKRPSPNLTSPFLGCOLA 951
Db 854 YPGILVNSKFKLVKDGQYMAQAPAPKNYSIMACWALETRAPTRPTQQICSLI--OKQ 911
Qy 952 EABEACIRTSIHLPKQAAPQ-----RGLRAQSPQRQ 984
Db 912 AQEDRRVPNYTNLPSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS 953

RESULT 5
TVMVMD
protein-tyrosine kinase (EC 2.7.1.112) fms precursor - feline sarcoma virus (strain McDo
A:Note: host Felis sp. (cat)
C:Date: 27-Nov-1985 #sequence_revision 31-Dec-1991 #text_change 13-Jun-1997
C:Accession: A00654
F:Hampe, A.; Gobet, M.; Sherr, C.J.; Galibert, F.
Proc. Natl. Acad. Sci. U.S.A. 81, 85-89, 1984
A:Title: Nucleotide sequence of the feline retroviral oncogene v-fms shows unexpected ho
A:Reference number: A00654; MUID:84119469; PMID:6582485
A:Accession: A00654
A:Molecule type: DNA
A:Residues: 1-941 <HAM>
C:Comment: This protein is synthesized as a gag-fms polyprotein.
C:Genetics:
A:Gene: fms
C:Superfamily: macrophage colony-stimulating factor 1 receptor; immunoglobulin homology;
C:Keywords: ATP; autophosphorylation; glycoprotein; kinase-related transforming protein;
otein kinase
F:1-23/Domain: signal sequence #status predicted <SIG>
F:24-941/Product: protein-tyrosine kinase fms #status predicted <MAT>
F:24-509/Domain: extracellular #status predicted <EXT>
F:35-86/Domain: immunoglobulin homology <IMM1>
F:120-179/Domain: immunoglobulin homology <IMM2>

```

```

F:217-280/Domain: immunoglobulin homology <IMM3>
F:316-381/Domain: immunoglobulin homology <IMM4>
F:410-484/Domain: immunoglobulin homology <IMM5>
F:510-534/Domain: transmembrane #status predicted <TMM>
F:535-941/Domain: intracellular #status predicted <INT>
F:577-915/Domain: protein kinase homology <KIN>
F:585-593/Region: protein kinase ATP-binding motif
F:42-84,127-177,224-278,417-482/Disulfide bonds: #status predicted
F:45,73,94,153,275,286,302,335,410,477,490/Binding site: carbohydrate (Asn) (covalent)
F:613,630,776/Active site: Lys, Glu, Asp #status predicted
F:781,794/Binding site: magnesium (Asn, Asp) #status predicted

Query Match 23.7%; Score 1247; DB 1; Length 941;
Best Local Similarity 33.2%; Pred. NO. 1.1e-56;
Matches 325; Conservative 162; Mismatches 347; Indels 144; Gaps 31;

Qy 75 TVYRAATVEAEGSITLQVQLATPGDSLCL-----W---VFKHSSLCGCPHFQDLQNRG 125
Db 13 TAWHAQGVPIQPSGPELVVEPGITVTLRCVNGSVGEWDGPISPHNLDLDPSSI----- 68
Qy 126 IVSMAILNVTTQAGEYLLHI---QSERANYVTLFTVNVRTQLYVL---RRPY----- 173
Db 69 ---LTTNATFQNTGY---HCTEPGNPRGNATI-----HLVYKDPAPKVKLAQE 114
Qy 174 FRKMNQDALL-CISEGVPEPTVEVLCSHRESCKEESGPAVRKEEKVHLBELFGTDI-- 230
Db 115 VTVLEGQDALLPCL---LTDPALE-----AGSVLRVVRGRPVLRQNTNYSFSPWHGFTIHK 166
Qy 231 -----RCARNALGRECTKL-----FIDLNQAPQSTL--POLFLKV-CEPLWIRCK 274
Db 167 AKFIENHYQCSAR--VDGRVTVMGIMLVKQKDISGPATLTLEPAELVRIQEAQAIVCS 225
Qy 275 AIHVNHGFLWELEDKALEEGSVFEMSTYSTNRMIRILLAFVSSVGRNDTGYTCSSS 334
Db 226 ASNDVNFVDSLRHGDTKL---TISQSDFDHNRVQ--KVLTLNLDHVSFQDAGNCTAT 281
Qy 335 K---HPSQASLVITILEKGFINATSQE--EYRIDPYEKFCFSVRKAYPRIIC-TWI--- 385
Db 282 NAWGNHSAFVRVVEAYLNLTSQSLLEQVTVGEKVDLQVKKVAYPGLESFNNYTLGP 341
Qy 386 FSQASFPCEORLENGYSKPCD-----HKNKPGYIFVAENDDAQFTQVFNIRKPKQ 441
Db 342 FSDYQDKLDFVTKYTRYSLSLPRKRSAGYSFLARNAGQONALTPELLTLYRYPE 401
Qy 442 VLANASQAS-----CSSDGYPLPSWTKKCDKSPNCTEE---IPEGVNNKANKRVFG 494
Db 402 VRVTMLINGSDTLLCEASGYQPSVTVVQCRSHTRDCDEAGLVLESHSEVLQSVPFH 461
Qy 495 QWSSSTLNMSAGKGLLVKCCAYNMGTSCTETFLNSPGPPFFIODNISFYATIGLC-- 552
Db 462 EVIVHSLAIGTLEHNRVTECFRNFVSGNSQTFWPISIGAHTQPLDELLFPVLLTCMS 521
Qy 553 -LPFTIVLVILCHYKQKQRYESQLOMIVTGLDNEYFYVDFRDYDLKWEPPRENIL 611
Db 522 IMALLLLLLLLLYKQKPKYQVRWKIE--SYEGNSYTFIDPTQLPNEKWEPPRNIL 579
Qy 612 EFGKVLGSAGFRVNNATAYGISTGVSIOVAVKMLKEKADSCKEALMSELKMMTHLGH 671
Db 580 QFGKTLGAGAFGVVEATAFGLGKEDAVLKAVKMLKSTAHADKEALMSELKIMSHLQ 639
Qy 672 HDNIVNLGACTLSPGVVLIPEYCCYGDLLNLYRSKRE-----KF 711
Db 640 HENIVNLGACTHGGPVLIVTEYCCYGDLLNLYRSKRE-----KF 711
Qy 712 HRTWTEIFKEHNFSSYPTFOAHNSMPCSRVQLHPPLDQSGFNGSIHSEDEIYEN 771
Db 700 HLEKYYVRDSDFS-----SQVDYTVEMRPVSTSSNDSEEDL----- 740
Qy 772 QKRLAEBEEDNLVITFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLTGHKVKIC 831
Db 741 -----GKEDGRPLELRLDLHFSQVAGQMAFLASKNCIHRDVAARNVLTSGRVAKIG 793
Qy 832 DFLGLARDILSDSSVYVRGNARLPVKWAPESLFEIGIYIKSDVWSYGILLWEIFSLGVP 891
Db 794 DFLGLARDILSDSSVYVRGNARLPVKWAPESLFEIGIYIKSDVWSYGILLWEIFSLGVP 853
Qy 892 YPGIPVDANFYKLIQSGFKMEOPFVATGIVFMQSCWAFSRKRPSPNLTSPFLGCOLA 951
Db 854 YPGILVNSKFKLVKDGQYMAQAPAPKNYSIMACWALETRAPTRPTQQICSLI--OKQ 911
Qy 952 EABEACIRTSIHLPKQAAPQ-----RGLRAQSPQRQ 984
Db 912 AQEDRRVPNYTNLPSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS 953

```

```
Db 794 DFGLARDIMDSNVIVKGNARLPVKWMAPEISFDCVYTVQSDVMSYGILLWEISLGLNP 853
Qy 892 YGPIVDANFYKLTQSOFKMEQPYATGIIYVWQSWAFDSRKRPPSPNLTSLFGCQLA 951
Db 854 YPGILVNSKFYKLVKDGQYMAQPAFAPKNIYISIMQACWALEPRTTRPTFQQICSL--QKQ 911
Qy 952 EAEACIARTSIHLPKQAA 969
Db 912 AQEDRRVPNTNLPSSSS 929

RESULT 6
A49814
protein-tyrosine kinase (EC 2.7.1.112), receptor type kit precursor - rat
C/Species: Rattus norvegicus (Norway rat)
C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 16-Jun-2000
C/Accession: A49814; S49088
R/Tsujimura, T.; Hirota, S.; Nomura, S.; Nomura, M.; Tono, T.; Morii, E.; Kikuchi, T.
Blood 78, 1942-1946, 1991
A/Title: Characterization of Ws mutant allele of rats: a 12-base deletion in tyrosine kinase gene
A/Reference number: A49814; MUID:92003944; PMID:1912577
A/Accession: A49814
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-978 <TSU>
A/Cross-references: GB:D12524; NID:g220707; PIDN:BA02094.1; PID:g220708
R/Tsujimura, T.; Tono, T.; Yamazaki, M.; Nomura, S.; Kitamura, Y.
submitted to the EMBL Data Library, October 1991
A/Description: Two isoforms of rat c-kit receptor tyrosine kinase.
A/Reference number: S49088
A/Accession: S49088
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-511,516-978 <TS2>
A/Cross-references: EMBL:X62491; NID:g509135; PIDN:CAA44354.1; PID:g509136
C/Superfamily: macrophage colony-stimulating factor 1 receptor; immunoglobulin homology;
C/Keywords: ATP; immunoglobulin homology <IM>
F/129-188/Domain: immunoglobulin homology <IM>
F/589-932/Domain: protein kinase homology <KIN>
F/597-605/Region: protein kinase ATP-binding motif

Query Match 23.4%; Score 1232.5; DB 1; Length 978;
Best Local Similarity 32.2%; Pred. No. 6.3e-56;
Matches 323; Conservative 176; Mismatches 340; Indels 165; Gaps 37;

Qy 36 CVLISHENNGSSAGKPSYRMVGRSPEDLQCTPRQSEGTVEATVEVAESGSIITLQVQ 95
Db 12 CVLLVLRGQTGTSQPSA-----SPGEPPSIQPAQSELIVRAGD-TIRLT 57

Qy 96 LATPGDLSCLVFKHSLGCGQPHDLQNRGIVSMALNV-TETQAGEYLLHIQSERANYT 154
Db 58 CTDPAFVK--WTFP-----ILDVRIENKQSEWIR--EKAEATHT 92

Qy 155 VLFT-----VNVRD--TQLYVLRPFRKMNQDALLCISGVPEPTV-EWVLC 200
Db 93 GKTYCVSGSLGRSIYFVRDPAVLFLVGLFLGKEDNDALVRC--PLTDPQVSNYSLI 149

Qy 201 SSHRESC-----KEGPAVVRKEKVLHFGTDIRCCARNALGRECTKLFTIDNQ 252
Db 150 ECDGKSLPTDLKFPVNPFGATITKVRVHRLC---IRCAQREGKWRMSDKFTLVRA 206

Qy 253 A-----POSTLPQL--FLKVGEPMLRCKAIHNVHFGLTW-----ELEDKALEE 295
Db 207 AIKAIPVVSVPETSHLLKEGDTFTVICTIKDVSIVSDMWIKLNPQSQKAVKRNWHQ 266

Qy 296 GSTFEMSTYTNRTMIRILLAFVSGVGRNDTGYTCSKSHPSQALVT---ILEKGFIN 352
Db 267 GDF-----NYERQETLT-----ISSARVDSGVFMCYANNNTFGSANTVTTLKVVKEGFIN 316

Qy 353 --ATSSQEEYEIDPYEKFCSVRKAYR-IRCTWIFSQASFPCEQGLE-----DGYVIS 405
Db 317 IFPVKNTTVFTDG-ENVDLVVEPEAYPKPHQOQIYNNRT--PTNRGEDIYKSDNQSN 373
```

```
Qy 406 KPCDH-----KNKPGEYIFYAENDDAQFTMTFLNIRKKPOVLA--NASASQASCSD 456
Db 374 RYVNELRLRLKGTGEGTYTFLVNSDVSAASVTFDVYVNTKPEILTYDRLMNGRLQCVAA 433
Qy 457 GYPLSPWTWKCKSDKSPNCTEEIPE-GVWNKANKRVFGOWVSSSTLNNSEAGKLLVKC 515
Db 434 GFPEPTIDWFTCTGABQRCITVPVPDVOIQNASVSPGKLVVQSSIDSSVFRHNGTVEC 493
Qy 516 CAYNSMTGSCBTIFLNSPGPPF-----IQDN-----ISFYATIGLCPLPFIWLIV 561
Db 494 KASNAVGKS--SAFFN---PAFKGNSKEQIQPHLFTPELLIGFVVTAGL-----NGIIVM 543
Qy 562 LICHKYKQFRYSOLOMI-QVTGPDNEYFVDFRDYEDLKWEPFRENLEFGKVLGSG 620
Db 544 VLAYKYLQKPMYEQWKVVEING---NNYVYIDPTQLFYDHWKWEFFPNRLSFGKTLGAG 600
Qy 621 AFGVMMATAYGISKTGVSIOVAVKMLKEKADSCKEALMSELKOWMTHLGHHDHNVNLLG 680
Db 601 AFGKVEATAYGLIKSDAAMTVAVMLKPSAHLTEREALMSELKVLSTYLGHNHNVNLLG 660
Qy 681 ACTLSGVPYLIIFYCCYCGDLLNLYRSKREKPHRTWTETFEKHNFSSTSYPTFOAHSSNMPG 740
Db 661 ACTVGGFTLVITEYCCYCGDLLNFLRRKDSF--IFSKQEEQADAAALYKNLLHSKSSCDS 718
Qy 741 SREVQLHPPLDQJSGFN-----GNSIHSDEIEYENOKRLAEDEEDLNVLTFBDL 791
Db 719 SNEY-----MDMKPGVSVVPTKTKRSARIDSYIERDVTPTAIMEDELALDL---BDL 770
Qy 792 LCPAYOVAKGMEFLFPKSCVHRDLAARNVLTHGVKVKICDFGLARDILSSSYVVRGNA 851
Db 771 LSFYQVAKGMAFKASKNCHRDLAARNILTHGRITKICDFGLARDIRNDSNYVYVKGNA 830
Qy 852 RLPVKWMAPESLPBGIVTIKSDVMSYGILLWEISLGVNVPYGPVVDANFYKLIQSGFKM 911
Db 831 RLPVKWMAPEISFNCVTVTFSDVMSYGIFLWELFSLGSSPYGMPVDSKFTYMKIEGFRM 890
Qy 912 EQPYATEGIYVWQSWAFDSRKRPPNLTSLFGCQLAABE 955
Db 891 LSPEHAPAAWYEVNKTCDADPLKRPFTKQVQVLIQKISDSSK 934

RESULT 7
TVMSKT
protein-tyrosine kinase (EC 2.7.1.112), receptor type kit precursor - mouse
N/Alternate names: tyrosine kinase receptor c-kit
C/Species: Mus musculus (house mouse)
C/Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 31-Mar-2000
C/Accession: S00474; B44876; I49596
R/Qiu, F.; Ray, P.; Brown, K.; Barker, P.E.; Jhanwar, S.; Ruddle, F.H.; Besmer, P.
EMBO J. 7, 1003-1011, 1988
A/Title: Primary structure of c-kit: relationship with the CSP-1/PDGF receptor kinase f
A/Reference number: S00474; MUID:88296403; PMID:2456920
A/Accession: S00474
A/Molecule type: mRNA
A/Residues: 1-975 <QIU>
A/Cross-references: GB:Y00864; NID:g50423; PIDN:CAA68772.1; PID:g50424
R/Rossi, P.; Marziani, G.; Albanesi, C.; Charlesworth, A.; Ghermà, R.; Sorrentino, V.
Dev. Biol. 152, 203-207, 1992
A/Title: A novel c-kit transcript, potentially encoding a truncated receptor, originate
A/Reference number: A44876; MUID:92331813; PMID:1378413
A/Accession: B44876
A/Molecule type: DNA
A/Residues: 771-814 <ROS>
A/Note: sequence extracted from NCBI backbone (NCBIN:108837, NCBI:108840)
R/Yasuda, H.; Galli, S.J.; Gelsler, E.N.
Biochem. Biophys. Res. Commun. 191, 893-901, 1993
A/Title: Cloning and functional analysis of the mouse c-kit promoter.
A/Reference number: I49596; MUID:93221533; PMID:7682073
A/Accession: I49596
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-22 <RES>
```

A;Cross-references: GB:L11358; NID:g293325; PIDN:AAA37420.1; PID:g293326
 C;Genetics:
 A;Gene: kit; C-kit
 A;Map position: 5
 C;Superfamily: macrophage colony-stimulating factor 1 receptor; immunoglobulin homology;
 C;Keywords: ATP; autophosphorylation; glycoprotein; kinase-related transforming protein;
 fic protein kinase
 F;1-22/Domain: signal sequence #status predicted <SIG>
 F;23-975/Domain: protein-tyrosine kinase kit #status predicted <KTC>
 F;23-975/Domain: extracellular #status predicted <EXT>
 F;51-100/Domain: immunoglobulin homology <IMM1>
 F;130-189/Domain: immunoglobulin homology <IMM2>
 F;221-235/Domain: immunoglobulin homology <IMM3>
 F;331-397/Domain: immunoglobulin homology <IMM4>
 F;426-496/Domain: immunoglobulin homology <IMM5>
 F;520-542/Domain: transmembrane #status predicted <TM>
 F;543-975/Domain: intracellular #status predicted <INT>
 F;586-929/Domain: protein kinase homology <KIN>
 F;594-602/Region: protein kinase ATP-binding motif
 F;598-98,137-187,234-293,431-494/Disulfide bonds: #status predicted
 F;146,296,303,323,355,370,466,489/Binding site: carbohydrate (Asn) (covalent) #status pr
 F;622,639,790/Active site: Lys, Glu, Asp #status predicted
 F;795,808/Binding site: magnesium (Asn, Asp) #status predicted

Query Match 23.2%; Score 1223; DB 1; Length 975;
 Best Local Similarity 32.1%; Pred. No. 1.9e-55;
 Matches 324; Conservative 172; Mismatches 331; Indels 182; Gaps 39;

Qy 36 CVLISHENNGSAGKPSYVRVGRSGPEDLOCTPRQSGTYVEATVEAESGSLTQVQ 95
 Db 12 CVLIVLLRGQATSPQSA-----SPGEFSPSTHPAOSLIVEAGD-TL--- 54
 Qy 96 LATPGDLISCL-----WVFKHSLGQCQPHF-----DLQNRGVISMAILNVTQAGEYLL 144
 Db 55 -----SLTCIDPDFVRWTFK-----TYFEMVENKKNWLOEKA-----EATRTGTYTC 98
 Qy 145 HQISERANYTLVFNVRD-TQLYVLRPRYFRKMNQDALL-----CIS 187
 Db 99 ---SNSNGLTSSIIYFVRDPAKFLVGLPLFGK-EDSDALVRCPLTDQVSNYSLEICDG 154
 Qy 188 EG-----VPEPTVEWLCSHRESCKEKGPAVRKEKVLHFGTDIRCCARNALGR 240
 Db 155 KSLPTDLTFVNP-----KAGITIKNVRAYH-----RLCVCAQOR 191
 Qy 241 ECTKL-----FTIDLNQA-----POSTLPOL--PLKVGELPWRCKAIHNGHGLTW---- 286
 Db 192 DGTWLHSDKFTLVKREAIKAPVSVPTSHLLKGGDTFTVVCTIKDVSTSVNSWLKN 251
 Qy 287 -----ELEDKALEEGSYFEMSTYSTRNIRILLAFVSSVGRNDGTGYTCSSSKHPQS 339
 Db 252 PQPHIAQVKHNSWHRGDF---NYERQETLT-----ISSARVDDSGVFNVCYANNFTGS 301
 Qy 340 SALVT---ILEKGFNATS-SQEEYIDPYEKFCFSVRPKAYPR-IRCTWTFPSQASPCPE 394
 Db 302 ANVTTLTKVVEKGFNINSPVKNNTTVFVTDGENVDLVVEYAYPKPEHQOQWIMRNT--SA 359
 Qy 395 ORGLE-----DGVYSIKFCDH-----KNKGEYIFYAENDDAQPTKMTFLNIRKKPOVL 443
 Db 360 NKGKDYKVDKNSNIRYVQNLRLRLKTEGTYTFLVNSDASASVTNFTVNTKPEIL 419
 Qy 444 A--NASASQASCSDDGYPLPWTWKKCKSPNCTEEI-PEGVWNKKNRKFVQGWVSSS 500
 Db 420 TYDRLLINGMLQVCAEGPPEPTIDWYFCTGAQRCTTPVSDVQVQVNSVSPFGKLVVQS 479
 Qy 501 TLNMSKAGLLVKKCAVNSGTSCETIFLNSPGFPFIQNIISYATIGLCLPIV--- 557
 Db 480 SIDSSVFRHNGTVECKASNDVGKS--SAFFN-----FAFKEQIQANTLTFTPLLIGFWAAG 533
 Qy 558 ---VLIVLICHYKKOPRYESQLOMI-QVTGELDNEYFVDFRDVEYDLKWEFFRENLEF 613
 Db 534 AMGIIVMWLTLYKLOKMYEYQVQWKEEING---NNYIIDPTQLPYDHKWEFFRNRUSF 590
 Qy 614 .GKVLGSGAFGRVMNATAYGISKTGVSIGQAVKMLKEKADSCKEALMSLKMTHLGHHD 673

RESULT 8

TVHUND

N;Contains: protein-tyrosine kinase (EC 2.7.1.112) csflr/fms
 C;Species: Homo sapiens (man)
 C;Date: 28-Dec-1987 #sequence revision 31-Dec-1993 #text change 16-Jul-1999
 C;Accession: S08123; A24533; I56672; I57648; I59083; I52772
 R;Hampe, A.; Shamooin, B.M.; Gobet, M.; Sherr, C.J.; Galibert, F.
 Oncogene Res. 4, 9-17, 1989
 A;Title: Nucleotide sequence and structural organization of the human FMS proto-oncogen
 A;Reference number: S08123; MUID:89239490; PMID:2524025
 A;Accession: S08123
 A;Status: nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA
 A;Residues: 1-972 <HAM>
 A;Cross-references: GB:U63963; EMBL:X14720; NID:g1915975; PIDN:AAB51696.1; PID:g1915976
 A;Note: this sequence was submitted to the EMBL Data Library, March 1989
 R;Cousens, L.; Van Beveren, C.; Smith, D.; Chen, E.; Mitchell, R.L.; Isaac, C.M.; Ver
 Nature 320, 277-280, 1986
 A;Title: Structural alteration of viral homologue of receptor proto-oncogene fms at car
 A;Reference number: A24533; MUID:86175013; PMID:2421165
 A;Accession: A24533
 A;Molecule type: mRNA
 A;Residues: 1-53, 'A', '55-972 <COU>
 A;Cross-references: GB:J03149
 A;Note: the authors translated the codon GCA for residue 54 as Pro
 R;Wheeler, B.F.; Roussel, M.F.; Hampe, A.; Walker, M.H.; Fried, V.A.; Look, A.T.; Rette
 J. Virol. 59, 224-233, 1986
 A;Title: The amino-terminal domain of the v-fms oncogene product includes a functional
 sequences.
 A;Reference number: I56672; MUID:86281820; PMID:3525854
 A;Accession: I56672
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: DNA
 A;Residues: 1-16 <RES>
 A;Cross-references: GB:M14002; NID:g182676; PIDN:AAA35849.1; PID:g553292
 R;Visvader, J.; Verma, I.M.
 Mol. Cell. Biol. 9, 1336-1341, 1989
 A;Title: Differential transcription of exon 1 of the human c-fms gene in placental trop
 A;Reference number: I57648; MUID:89261741; PMID:2524648
 A;Accession: I57648
 A;Status: preliminary
 A;Molecule type: mRNA
 A;Residues: 1-16 <RES>
 A;Cross-references: GB:M25786; NID:g349454; PIDN:AAA58421.1; PID:g553224
 R;Browning, P.J.; Bunn, H.F.; Cline, A.; Shuman, M.; Nienhuis, A.W.
 Proc. Natl. Acad. Sci. U.S.A. 83, 7800-7804, 1986
 A;Title: Replacement of COOH-terminal truncation of v-fms with c-fms sequences marked
 A;Reference number: I59083; MUID:87017034; PMID:3532121

A:Accession: I59083
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 874-972 <R3>
A:Cross-references: GB:M14193; NID:g182521; PIDN:AAA35834.1; PID:g182522
R:Nienhuis, A.W.; Bunn, H.F.; Turner, P.H.; Gopal, T.V.; Nash, W.G.; O'Brien, S.
Cell 42, 421-428, 1985
A:Title: Expression of the human c-fms proto-oncogene in hematopoietic cells and its del
A:Reference number: I52772; MUID:85282599; PMID:4028159
A:Accession: I52772
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 244-295 <R4>
A:Cross-references: GB:M11067; NID:g182674; PIDN:AAA35848.1; PID:g442423
C:Genetics:
A:Gene: GDB:CSF1R; FMS
A:Cross-references: GDB:120600; OMIM:164770
A:Map position: 5q33.2-5q33.3
A:Introns: 17/1; 103/1; 198/1; 243/3; 297/1; 361/2; 400/1; 440/2; 504/1; 542/3; 585/1; 6
C:Superfamily: macrophage colony-stimulating factor 1 receptor; immunoglobulin homology;
C:Keywords: ATP; autophosphorylation; glycoprotein; kinase-related transforming protein;
fic protein kinase
F:1-23/Domain: signal sequence #status predicted <SIG>
F:24-972/Product: macrophage colony-stimulating factor 1 receptor #status predicted <MA
F:24-512/Domain: extracellular #status predicted <EXT>
F:35-86/Domain: immunoglobulin homology <IMM1>
F:120-179/Domain: immunoglobulin homology <IMM2>
F:217-280/Domain: immunoglobulin homology <IMM3>
F:316-383/Domain: immunoglobulin homology <IMM4>
F:412-487/Domain: immunoglobulin homology <IMM5>
F:513-537/Domain: transmembrane #status predicted <TM>
F:538-972/Domain: intracellular #status predicted <INT>
F:580-917/Domain: protein kinase ATP-binding motif
F:588-596/Region: protein kinase ATP-binding motif
F:42-84,127-177,224-278,419-485/Disulfide bonds: #status predicted
F:45,73,153,240,275,302,353,412,428,480/Binding site: carbohydrate (Asn) (covalent)
F:616,633,778/Active site: Lys, Glu, Asp #status predicted
F:763,796/Binding site: magnesium (Asn, Asp) #status predicted

Query Match 23.2%; Score 1221.5; DB 1; Length 972;
Best Local Similarity 32.5%; Pred. No. 2.3e-55;
Matches 330; Conservative 152; Mismatches 331; Indels 201; Gaps 33;

QY 76 VYEAATVEAEGSITLQVOLATPGDLSCLVFKHSLGQPHDQNRGIVSMAIL--- 132
DB 22 VIEPSPELVVKGAT--VTLRGVNGSVEMDGPPS-----PHWTLYSDG--SSSILSTN 72

QY 133 NVTETQAGEYLLHIQSERANVTVLTNVVDRDTQLYVLRPY-----FRKMENODALL-C 185
DB 73 NATFQNTGTYRCTEPGDPGLGSAIHLVYKDP-----ARPWNVLAEVVFVDDQALLPC 127

QY 186 ISGVPEPTVEVWVLCSSHRESCKEKGPAVVRKEK-----VLHE---LF 226
DB 128 L---LTDPLV-----EAGSVLRVGRPLMRHTNYSFSPWHGFTIHRAKFIQ 171

QY 227 GTDIRCCARNALGRECTKLTIDLNOAPOSTLP-----QLFKVGEPLWIRKAI 276
DB 172 SDQYQCSALMG-GR---KVMSSIRLVQKVIQPGPALTLPVPAELVIRGEAAQIVCSAS 227

QY 277 HVNHRGFLTWELKEDKALEGSEYFEMSTYSTNRTMIRILLAFVSVGRNDTGYTCSS-- 334
DB 228 SVDVNFDFVLOHNTKL---AIFQSDFNHRVQ-KVLTNLNQDVPQHAGNYSVCASNV 283

QY 335 --KHPQSALVTILEKGFINATSSQE--EYBIDPYEKFCFVRPKAYPRIR--CTWIFSQAS 390
DB 284 QGKH-STSMFRVRESAYLNLSEQNLIQEVTVGEGNLKVMVEAYVGLQGFNWTY---- 338

QY 391 FPCQRGLEDGYSIKFCDHNKP-----GYIFVAEN 423
DB 339 -----LGPFSHQPEPKLANATTQYRHTFTLSLPLKPSBAGRYSFLARN 385

QY 424 DDAQFTKMTFLNKKPKQV-----LANASASQACSSDGYPLPSWTKKSDKSPNCTEE 478

DB 386 PGWRALTFTLTLRYPPEVSVIWTPIFNGSGT-LLCAASGYPPQPNVTWLCQSGHTDRCDRA 444
QY 479 IPEGVWN-----KXANKRVFGQWVSSSTLNMSEAGKGLLVKCCAYNSMGTSCTETFLNSPG 534
DB 445 QVLQWDDPYEVLVSQEPHFVTVQSLTTLVETLEHNTQVECKRAHNSVSGSWAFIPISAG 504
QY 535 PPFQIDNISFYATIGLC--LPFIVLVILVILCHYKQKQFYESOLOMIQVITGPDLDNEYF 591
DB 505 AHTHPDDEFLTPVVVACMSIMALLLLLLLYKQKPKYQVRWKIIE--SYEGNSVT 562
QY 592 YVDFRDYEDLKBFPFRENLEFGVLGSGAFGRVVMNATAGISKTGVSIOVAVKMLKEKA 651
DB 563 FIDTQLPYNEKNEFPNNLQFGKTLGAGAFGKVVEATAFGLGKEDAVLKVAVRMLKSTA 622
QY 652 DSCKEALMSLKMVTHLGHHDNINVLGACTLGGPVYLIFFEYCCYGLLNLVLRKREKF 711
DB 623 HADEKEALMSLKMVTHLGHHDNINVLGACTLGGPVYLIFFEYCCYGLLNLVLRKREKF 680
QY 712 HRTWTEIFKEHNFSSYTFQAHSSNMPGSRVQLHPPLDQLSGNPGNSIHSEDE---- 766
DB 681 -----AMLGP---SLSPQDPEGVDYKNIHLEKYYRRD 712
QY 767 -----IYENQKELABE--EEDNLVLTFFEDLLCLCFAYQVAKGMEFLEPK 808
DB 713 SGFSSQGVDTYVEMRPVSTSSNDSFSEQDLKDGRLPLELRLDLHFSSQVAGMAFLASK 772
QY 809 SCVHRDLAARNVLVTHGKWKVICDFGLARDLSDSSYVVRGNARLPVKNAPESLPEGIY 868
DB 773 NCITHRVAARNVLTNGHVAKIGDFGLARDIMDSNVIVKGNARLPVKNAPESLFDVCY 832
QY 869 TIKSDVMSYGILLWEIFSLGVNYPGIPVDANFYKLIQSGFKMBOQPFYATEGIYFVMQSC 928
DB 833 TVQSDVMSYGILLWEIFSLGVNYPGIPVDANFYKLIQSGFKMBOQPFYATEGIYFVMQSC 892
QY 929 WAFDSRKRPSFPNLTSLGQLAEAEACRTSTHLPKQAAPOQRGGLRAOSQPQ 982
DB 893 WALEPHTHRPTFOQICSLF--QEQAQEDRRERDYNLFPSS---RSGGSGSSSE 941

RESULT 9
I51703
C-kit-related kinase 1 (XKrk1) - African clawed frog
C:Species: Xenopus laevis (African clawed frog)
C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 18-Jun-1999
C:Accession: I51703
R:Baker, C.V.; Sharpe, C.R.; Torpey, N.P.; Heasman, J.; Wylie, C.C.
Mech. Dev. 50, 217-228, 1995
A:Title: A Xenopus c-kit-related receptor tyrosine kinase expressed in migrating stem c
A:Reference number: I51703; MUID:95344996; PMID:7619732
A:Accession: I51703
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-954 <BAK>
A:Cross-references: EMBL:Z48770; NID:g763033; PIDN:CAA8688.1; PID:g763034
C:Superfamily: macrophage colony-stimulating factor 1 receptor; immunoglobulin homology
C:Keywords: ATP
F:575-915/Domain: protein kinase homology <KIN>
F:583-591/Region: protein kinase ATP-binding motif

Query Match 23.2%; Score 1220; DB 2; Length 954;
Best Local Similarity 32.0%; Pred. No. 2.7e-55;
Matches 315; Conservative 160; Mismatches 321; Indels 188; Gaps 35;

QY 77 YEAATVEAEGSITLQVOLATPGDLSCLVFKHSLGQPHDQNRGIVSMAIL--- 130
DB 15 YTDGVKINDGDRVTNVGVKVSLECRDAHLVTLAFQKSLGMLKKPR-DLKSRLP----- 68

QY 131 ILNVET-----QAGEYLHIQSERANVTVLTNVVDRDTQLYVLRPYRPFKME 178
DB 69 -LNNSETDQFFVIKADLPHIGYICTNTETQNTSV--SLFVKDPAFPDLPDIFDVT 125

QY 179 NODAL-LCISEGVPEPTVEVWVLCSSHRESCKEKGPAVVRKEKVLHLEFDTDIRC----- 232

Db 126 GADTVGMCFF--TDPMDIAI-----EKC--DGSPLPE-----NFTFTDIEAGTIK 169
QY 233 -----CARNALGR-ECTKLTFTIDLNQAPQSTLPOLFL-----KVGEPLMIRC 273
Db 170 TVQLAFDSCVVCNGSKSGTGVKSSSTSIHVKVPKK-VPIVFLSKSQPLVKTGEPPEVTC 228
QY 274 KAIHVNHGFLTW-ELEDKALEBGSYFEMSTYSTNRTMIRILLAFVSSVGRNDTGYTCS 332
Db 229 AVLVDVSTVKAQWLVDKGVTKQANFRSSNVFSLTKS-----DGVPSSESRFTQC 282
QY 333 SSKHPQO---SALVTILEKGFINATSQE-EYEDPVEKECFSVRFKAYPR-IRCTWIFS 387
Db 283 AENAIGQVNAFTFLDVIDGVNLTVENTTISVAGDNILVKYIDAYPHPDGVWYTF 342
QY 388 QASFPCEQRLGEGYSISKCDHKN-----KPGEIFYFAENDDAQFTKM 432
Db 343 NETL----LNTSHYATK--DEGNRYVSELHLRLKGTGKYVYFTYTTNSDDASVSP 396
QY 433 TLMIRKKPOVLANASQAS--CSSDGYPLPSWTWKCKSDKSNCTEIEIEGVWNKK--A 488
Db 397 NIQVKTREPEILAIERTSEGLQCVATGFPVPAIQWYFCPGSEORCTDYPPLSPVNEKFIQ 456
QY 489 NRKVFQWYSSSTLNSEAGKLLVKCCAYNSMGTS-----CETIFLNSPGP 535
Db 457 ENSSLGRIVVESTIDVNDLKKNGTVQCVASNEVESYVFSFAIKEKLRTHTLFT----- 511
QY 536 FPIQDNISFYATIGLCPPIVVLIVLICHKKQFQRYESQLOMI-OVTGPDONEYPVD 594
Db 512 -PLL---IGFIAAGL---MCTAVAVLMYKQPKYEQWKVVEING---NNVYIID 560
QY 595 FRDYEDLKWEPRENLEFGKVLGSGAFGRVMNATAYGISKTGVSIOVAVKMLKEKADSC 654
Db 561 PTQLPYDNKWEFPRDLFCGKILGAGAFGVVEATAYGLLKEDSRLTVAVKMLKPSAHS 620
QY 655 EKEALMSELKQWTHLGHHDNIVNLLGACTLSGPVLYFFCYCGDGLLNLYRSKREKPHRT 714
Db 621 EREALMSELKVLGYLGHKHNIVNLLGACTVGGTPLVTITEYCCYGDGLLNLYRKRDSF--- 677
QY 715 WTEIFEKHNFSYPTFOAHNSNMPGSRVQLHPPLDQLSGFNGNSIHSDEIEYENOK- 773
Db 678 -----ICPKFEDNS-----EALYKQL-----LNRDMGCGMGEYIDMKP 713
QY 774 -----RLABEEEDLNVLTFEDLLCFAYQVAKGMEFLEPKSC 810
Db 714 AVSYVVPVTKDKRRSGSFGQDVSVSIPEDDL-ALOTEDLINFYSYQVAGMFLASKNC 772
QY 811 VHRDLAARNVLVTHGKVKVVICDFGLARDILSDSSVYVGRNARLPVKWMAPESEFEGYTI 870
Db 773 IHRDLAARNLILLTHGRITKICDFGLARDIRNDSNYVYVGNARLPVKWMAPESEFHCYTF 832
QY 871 KSDVWSYGILLWEIFSLGVNYPGIPVDANFYKLIQSFGKMEOPFYATEGIFVMOQSCWA 930
Db 833 ESDVWSYGILLWEIFSLGVNYPGIPVDANFYKLIQSFGKMEOPFYATEGIFVMOQSCWA 930
QY 931 FDSRKRPSPFNLSFLGCOLAEAE 954
Db 893 SDPLKRPFTKQIVQMVQEQQLSDSK 916
RESULT 10
I45877
protein-tyrosine kinase (EC 2.7.1.112), receptor type kit precursor - aurochs
C:Species: Bos primigenius (aurochs)
C:Date: 19-Dec-1997 #sequence_revision 19-Dec-1997 #text_change 20-Apr-2000
C:Accession: I45877
R:Kubota, T.; Hikonou, H.; Sasaki, E.; Sakurai, M.
Gene 141, 305-306, 1994
A:Title: Sequence of a bovine c-kit proto-oncogene cDNA.
A:Reference number: I45877; MUID:94215924; PMID:7512939
A:Accession: I45877
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-977 <KUB>

A;Cross-references: GB:D16680; NID:G516659; PIDN:BAA04084.1; PID:G516660
C;Superfamily: macrophage colony-stimulating factor 1 receptor; immunoglobulin homology;
C;Keywords: ATP; phosphotransferase; tyrosine-specific protein kinase
F;329-395/Domain: immunoglobulin homology <IMM>
F;598-932/Domain: protein kinase homology <KIN>

Query Match 23.1%; Score 1215; DB 2; Length 977;

Best Local Similarity 32.8%; Pred. No. 5e-55;

Matches 304; Conservative 141; Mismatches 298; Indels 184; Gaps 28;

QY 134 VTE-----TQAGEYLLHIQSERANTYVLTNNVDTQ-LYVLRBPYPFKMENQDALLCI-- 186

Db 83 ITEKAEATNGTNC---TNKGLSSSIYVFDPEKFLIDLIDPLYGKEENDTLVRCELT 139

QY 187 -----SEGVEPEP-TVVEWLCSSHRESCKEGPAVVRKEEKVLHFGTDIRC-- 232

Db 140 DPEVTNYSLTGCGGKPLPKDLTFV-----ADPKAGITIRNVKREYHRL-----CLH 185

QY 233 CARNALGRE-CTKLTFTIDLNQAPQS-----TLPQLFLKVGEPLEWIRCKAIHVNHGFLT 285

Db 186 CSANQRGKSMLSKKFTLKVRRAIKAVPVSVSKTSYLLREGEFPAVTCCLKIDVSSSSVDSM 245

QY 286 WELEDKALEBGSYFEMSTYSTNRTMIRILLAFVSSVGRNDTGYTCSSSKHPSPQSALVTI 345

Db 246 WIKENSQOTKAOQTKKNSWHQGDPSYLRQERLTITISSARVNDSGVFMCVANTFGSANVTIT 305

QY 346 LE---KGFINA-TSSQBEYEIDPVEKFCFSVRKAYPR-IRCTWIFSQAQPPCPQORGLD 400

Db 306 LEVVDGKFINI FPMNTTTFVNDGENVDLVVEAYEPKPHROWIYNRT----- 355

QY 401 GYSISKPCDHK-----NKPGEYIFYAENDDAQFTMFTLINRKK 439

Db 356 --STDKWDVYPKSEBNSIRYVNEHLHLTLKLTGEGGYTTFHVSNSDVNSSTFVNVYNTK 413

QY 440 PQVLANASA--SOASCSDGYPLPSWTWKCKSDKSNCTEIEI-PEGVWNKANRKFQGW 496

Db 414 PEILTHDRLVNGLQCVAAAGPEPTIDWYFCPGTEQRCVFPVGVDPVQIQNSSYSPEGL 473

QY 497 VSSSTLNMSAGKGLVKCCAYNSMGTS-----CETIFLNSPGPPEFI 539

Db 474 VYVSTIDDDSTFRKNGTVCEKRAYNDVGKSSASFPAPKGNKQBQIHAHTLFT-----PLL 527

QY 540 QDNISFYATIGLCPPIVVLIVLICHKKQFQRYESQLOMI-OVTGPDNEIFYVDFRDY 598

Db 528 ---IGFVIRAGLWCI FVMIL---TYKYLQKPMYEVQWKVVEING---NNVYIIDPTQL 577

QY 599 EYDLKWEFPRENLEFGKVLGSGAFGRVMNATAYGISKTGVSIOVAVKMLKEKADSCKEA 658

Db 578 PYDHKWEFPRENRLSFGKTLGAGAFGVVEATAYGLIKSDAAMTVAVKMLKPSAHLTEREA 637

QY 659 LMSLKMWTHLGHHDNIVNLLGACTLSGPVLYLFEYCCYGDGLLNLYRSKREKPHRTWTEI 718

Db 638 LMSLKVSLYGLNHNMIIVNLLGACTIGGPTLVITEYCCYGDGLLNFLRRKRDSF----- 690

QY 719 FKEHNFSYPTFOAHNSNMPGSRVQLHPPLDQLSGFNGNSIHSDEI-----EYENQ 772

Db 691 -----ICSKQEDHAEVALYK-----NLLHSEKSSCNDSTNEYMDM 725

QY 773 K-----RLABEEEDLN-----VLTFEDLLCFAYQVAKGMEFLE 806

Db 726 KPGVSVVPVTKDKRRSARSIGSVIERDVTPTAIMEDDELALDLEDLLSFSYQVAKGMFLA 785

QY 807 FKSCHVHDLAARNVLVTHGKVKVVICDFGLARDILSDSSVYVGRNARLPVKWMAPESEFEG 866

Db 786 SKNCIHRDLAARNLILLTHGRITKICDFGLARDIRNDSNYVYVGNARLPVKWMAPESEFNC 845

QY 867 IYTIKSDVWSYGILLWEIFSLGVNYPGIPVDANFYKLIQSFGKMEOPFYATEGIFVMOQ 926

Db 846 VYTFESDVWSYGIFLWELFSLGSPFPGMPVDSKFFYKMKIEGFMFLSEHPAEPAMYDIMK 905

QY 927 SCWAFDSRKRPSPFNLTSLGCOLAEAE 953

Db 906 TCWDADPLKRPFTKQIVQMLIEKQISES 932

RESULT 11
S16385
macrophage colony-stimulating factor 1 receptor precursor - rat
N:Contains: protein-tyrosine kinase (EC 2.7.1.112) CSP-IR
C:Species: Rattus norvegicus (Norway rat)
C>Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 02-Jun-2000
C:Accession: I60321; S16385
R:Borycki, A.G.; Guillier, M.; Leibovitch, M.P.; Leibovitch, S.A.
Growth Factors 6, 209-218, 1992
A>Title: Molecular cloning of CSF-1 receptor from rat myoblasts. Sequence analysis and
A:Reference number: I60321; MUID:93001225; PMID:1389227
A:Accession: I60321
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-978 <RES>
A:Cross-references: EMBL:X61479; NID:957543; PIDN:CAA43706.1; PID:G57544
A>Note: in Genbank entry RRCST1, release 113.0, the source is designated as Rattus rattus
A:Note: submitted to the EMBL Data Library, August 1991
C:Superfamily: macrophage colony-stimulating factor 1 receptor; immunoglobulin homology;
C:Keywords: ATP; autophosphorylation; glycoprotein; growth factor receptor; kinase-relat
protein; tyrosine-specific protein kinase
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-978/Product: macrophage colony-stimulating factor 1 receptor #status predicted <MAI
F:20-515/Domain: extracellular #status predicted <EXT>
F:35-86/Domain: immunoglobulin homology <IMM1>
F:120-179/Domain: immunoglobulin homology <IMM2>
F:217-280/Domain: immunoglobulin homology <IMM3>
F:316-381/Domain: immunoglobulin homology <IMM4>
F:410-485/Domain: immunoglobulin homology <IMM5>
F:516-535/Domain: transmembrane #status predicted <TMM>
F:536-978/Domain: intracellular #status predicted <INT>
F:578-915/Domain: protein kinase homology <KIN>
F:586-594/Region: protein kinase ATP-binding motif
F:42-84,127-177,224-278,417-483/Disulfide bonds: #status predicted
F:45,73,302,335,389,410,449,478,491/Binding site: carbohydrate (Asn) (covalent) #status
F:614,631,776/Active site: Lys, Glu, Asp #status predicted
F:781,794/Binding site: magnesium (Asn, Asp) #status predicted

Query Match 23.0%; Score 1210; DB 2; Length 978;
Best Local Similarity 33.5%; Pred. No. 8.9e-55;
Matches 318; Conservative 147; Mismatches 329; Indels 154; Gaps 31;

QY 78 EAATVEAEGSGITLOVQLATPGDLSCLVFKHSIGCPHFLQNRIGVSMALNVTT 137
DB 36 ETVTLRCVNSGVEMD-----GPSFYWTLDPSGSS-----TLTRNATFK 77
QY 138 QAGEY-LLHIQGERANYTVLFTVNRDITOLYV-----LRPFYRKFNQDALL-CIS 187
DB 78 NTGTYRCTELEDPMAGSTTI-----HLVYKDPAHSNWNLIAQEVTVVGEQEAFLPCL- 128
QY 188 EGVPEPTVWVLCSSHRESCKEKGPAVVRK-----EBKVLHFLGTDIRCC 233
DB 129 --ITDPALK-----DSVSLMREGGRQVLRKTYFFSAWGRFIRKAKVL-----DSNTYVC 177
QY 234 ARNALGRCTKL-FTIDLN-----QAPQSTL-PQLFLV-GEPLWIRCKAIHNVHGFGLTW 286
DB 178 KTWNGRESTGTGILKVNVRHPEPPQIKLEPSKLVIRGEAAQIVCSATNAEVEGFNVIL 237
QY 287 BLEDKALE-EGSYPEMSTYSTNRTWIRILLAFVSSVSGNDGYTC-SSSKHPQSALV 343
DB 238 KRGDTKLEIPLNSDFQDNYKKVRL-----SLNAVDPQDAGIYCSVASNDVGRTRATM 291
QY 344 T--ILEKFINATSSQE-EYEIDPYEKFCFSVRKAYPRIR-CTWIFSQAPSPCBORGLE 399
DB 292 NFQVVESAYNLNLSQSLQEVSVGDSLLTTHADAYPSIQHYNTYLGPFPE-DQRKLE 350
QY 400 -----DCYISIKFCDH--KNKPGYIFYAENDDAQFTKMTLIRKPKQVLANASASQ 450
DB 351 FITORAIYRYTFKFLNRVKAESAQGYFLMAQNKAGWNLNLTFLRYRYPPEVSVTWPVN 410
QY 451 AS-----CSSDGYPLPSWTWKKCSKSPNCTEIPGVWN-----KKNRKYFGQWVSSSTL 502

DB 411 GSDVLFCDVSGYQPSVTWMECRGHTDRCDQAALQVWNDTHPEVLSQKPFQKVIISQQL 470
QY 503 NMSEAGKGLVKCCAYNSMGTSCTETIFLNSPPFPFIQDNISFYATIGLC---LPFIVVL 559
DB 471 PIGTLKHNMTYFCKTHNSVGNSSQYFRAVLSQSQKLPDESFTPTVVVACMSVMSLLVLL 530
QY 560 IVLCHKYKQFRYESQLQMIQ-VTGPLDNEYFYVDPRDYEDYDLKWEFFRENLEFGKVLG 618
DB 531 LLLLLYKQKPKYQVRWKIIRYEG---NSYTFIDPTQLPYNEKWEFFRNLIQFKTILG 587
QY 619 SGAGRVNNAATYGLTGVTSIOVAVKMLKEKADCEKALMSELKMMTHLGHHDNIVNL 678
DB 588 AGAFGKVVETAFAFGKGEDAVLKVAVKMLKSTAHADAKEALMSKLSKINSHLQCHENIVNL 647
QY 679 LGACTLGGPVVLIPEYCCYGLLNLVLSKRE-----KPHRTWTETI 718
DB 648 LGATHGGPVLVITYECCYGLLNLVLSKRE-----KPHRTWTETI 707
QY 719 FKEHNFSSYPTFOAHNSNMPGSRVQLHPPDLQSLGFGNSIHSDEIEYENQKRLAE 778
DB 708 RDSGFS-----SQGVDTYVEMRPVSTSSDSFFKQD-----L 740
QY 779 EEDLNVLITPDLCPAYQVAKMEFLFKSCVHRDLAARNVLVTHGKVKICDFGLARD 838
DB 741 DKPSRPLELDLWDLHFSQVQAQNAFLASKNCIHRDVAARNVLLTSGHVAKIGDFGLARD 800
QY 839 ILSDSSVYVRGNARLPVKWMAPESLFGIYTIKSDVMSYGLLWEIFSLGVPYGPVVD 898
DB 801 IMDSNVVYVRGNARLPVKWMAPESLIYCVYTVQSDVMSYGLLWEIFSLGVPYGPVVD 860
QY 899 ANFYKLTQSGFKMEQPPYATGTYGVVMSQWAFSRKPSFNLTSFL 946
DB 861 NKFYKLVKQGYQMAQVPFAPKNIYSINQSCWDLPTRRPTFQQICFL 908

RESULT 12
TVMSMD
macrophage colony-stimulating factor 1 receptor precursor - mouse
N:Contains: protein-tyrosine kinase (EC 2.7.1.112) csf1r/fms
C:Species: Mus musculus (house mouse)
C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 02-Jun-2000
C:Accession: S01880
R:Rothwell, V.M.; Rohrschneider, L.R.
Oncogene Res. 1, 311-324, 1987
A>Title: Murine c-fms cDNA: cloning, sequence analysis and retroviral expression.
A:Reference number: S01880; MUID:88217329; PMID:2966922
A:Molecule type: mRNA
A:Residues: 1-976 <ROT>
A:Cross-references: EMBL:X06368
C:Genetics: fms
C:Superfamily: macrophage colony-stimulating factor 1 receptor; immunoglobulin homology;
C:Keywords: ATP; autophosphorylation; glycoprotein; kinase-related transforming protein;
fic protein kinase
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-976/Product: macrophage colony-stimulating factor 1 receptor #status predicted <MAI
F:20-515/Domain: extracellular #status predicted <EXT>
F:35-86/Domain: immunoglobulin homology <IMM1>
F:120-179/Domain: immunoglobulin homology <IMM2>
F:217-280/Domain: immunoglobulin homology <IMM3>
F:316-381/Domain: immunoglobulin homology <IMM4>
F:410-485/Domain: immunoglobulin homology <IMM5>
F:516-535/Domain: transmembrane #status predicted <TMM>
F:536-976/Domain: intracellular #status predicted <INT>
F:578-914/Domain: protein kinase homology <KIN>
F:586-594/Region: protein kinase ATP-binding motif
F:42-84,127-177,224-278,417-483/Disulfide bonds: #status predicted
F:45,73,302,335,389,410,449,478,491/Binding site: carbohydrate (Asn) (covalent) #status
F:614,631,776/Active site: Lys, Glu, Asp #status predicted
F:781,794/Binding site: magnesium (Asn, Asp) #status predicted

Query Match	22.8%; Score 1198.5; DB 1; Length 976;
Best Local Similarity	34.1%; Pred. NO. 3.5e-54;
Matches 317; Conservative 148; Mismatches 347; Indels 117; Gaps 31;	
QY	78 EAATVEAEGSITLQVLAATGDLISLWFKHSSLCGPHFDLQNRGIVSMALNVTTT 137
DB	36 ETVTLRCVSGSVEWD-----GPISPIWTLDPESFGS-----TLTTSNATFK 77
QY	138 QAGEV-LHIIQSERANYTLFTNVRDQLYV-----LRRPYFKVENODALL-CFS 187
DB	78 NTGTVRCTELEDPMAGSTTI-----HLYKDPAHSNWLLAQEVTVVGEQAVLPCL- 128
QY	188 EGVPEPTVEWLVCSHRESCKEKGPAVVRKEEVLHLEFGTDTR-----CCARNA 237
DB	129 --ITDPALK-----DSVLSMREGGRQVLRKTVVFPSPWEGSIIRKAKVLDNSNTYVCKTWV 181
QY	238 LGRECKTL-FTIDLN-----QAQSTI-PQLFLKV-GEPLWIRCKAHVNHGRLTWLELD 290
DB	182 NCRESTSTGILWKNVRHPEPPQIKLEPSKLVIRGEAAQIVCSATNAEFGFNVLKRGD 241
QY	291 KALE--EGSYFEMSTYSTNRTMIRILLAFVSSVGRNDTGYTC-SSSKHPSOSALVT--I 345
DB	242 TKLEIPLNSDFQDNYKKVRAL-----SLNAVDFQDAGIYSCVANDVGTTRATWNVQV 295
QY	346 LEKGFINATSSOE-EYEIDPYEKFCSVRFKAYPRIR-CTWIFSQASFCEQGLE---- 399
DB	296 VESAYLNLTSQSLQEVSGDSLILTVHADAYPSIOHYNWTVLGPFFE-DQKLEFITU 354
QY	400 ----DGYSISKFDCH--KNKPGGEYIFAENDDAQFTKMTFLNTRKKQOVLANASASQAS-- 452
DB	355 RAIYRYTFLFLNRVASEAGQVFLMAQNKAGNNLTFELTLRYPPSEVSTWMPVNGSDV 414
QY	453 ---CSSDGYPLPMTWKCKSDKSNCTEETPEGVVN---KKNRKYFGVGSVSSSTLNMSE 506
DB	415 LFCDSVSGYQPSVTWMECKGHDRCDEAQAHLHNDTHPEVLSQKPFDKVIOQSQLPICP 474
QY	507 AGKGLLVKVCAYNSMGTSCTETPLNSFGPFPPFIQDNISFYATIGLC---LPFTVILVILI 563
DB	475 LKHNTYFYCKTHNSVGNSSQYFRAVLSGQSKQLPDESFLTPVVVACWSVMSLLVLLLLL 534
QY	564 CHYKQKQFYESQLMIQ-VTGPLDNEYVYVDRDYEDLKWFFPRENLEFGKVLGSGAF 622
DB	535 LYKYKQKPYQYRWKIIERYEG---NSYTFIDPTQLPYNEKWEFPFRNQLQFGKTLGAGAF 591
QY	623 GRVMNATAGISKTGVSIOAVKMLKADSKCEKALMSLKMTHLGHNDHNVNLLGAC 682
DB	592 GKVVEATAFGLGKEDAVLKVAVMKLSATAHDEKALMSLKMTHLGHNDHNVNLLGAC 651
QY	683 TLLSGPVYLIFECYCGDNLNLYRSKREKFTHTWTEIFKEHNFSSYPTFOAHNSSSMPGSR 742
DB	652 THGGPVLVYVTECCYGDHLNFLRRKAEAMHGP-----SLSPGQDSEGDSSYKNTH 701
QY	743 EVQLHPPPLDQLSGFNNSIHSDEDEIYENQKRLAEEDBDLN-----VLTFFDLCLCFAYQ 797
DB	702 LEKKYVVRD--SGFSSQGVDTYVEMRPFVSTSSDSFPFKQDLDEKHSRPLELMDLHFFSQ 759
QY	798 VAKGHEFLFKSCVHRDLAARNVLTGHKKVVKICDFGLARDILSDSYVVRGNARLPVKW 857
DB	760 VAQGMFLASKNCIHRDVAARNVLLTSGHVAKIGDFGLARDINDNSNVVVKGNA-LPVKW 818
QY	858 MAPESLFEIGYITKSDVMSYGILLWEIFSLGVNVPYGPVDFANFYKLIQSGFKNWGPFFYA 917
DB	819 MAPESIFDCVITVQSDVMSYGILLWEIFSLGLNPNYPYGIHVNKFKYLVKDYQMAQPVFA 878
QY	918 TEGIVPMQSCWAFDSRKPSFPNLSFLL 946
DB	879 PKNYISIMQSCWLEPTRRPTFOQICFL 907

RESULT 13

THUKT

N;protein-tyrosine kinase (EC 2.7.1.112), receptor type kit precursor - human
N;Alternate names: mast/stem cell growth factor receptor; tyrosine kinase receptor c-kit

C:Species: Homo sapiens (man)
C>Date: 31-Mar-1991 #sequence revision 31-Mar-1991 #text change 21-Jul-2000
C:Accession: S01426; PC1015; A41815; B41815; I37548; I56954; I54336
R:Yarden, Y.; Kuang, W.J.; Yang-Peng, T.; Coussens, L.; Munemitsu, S.; Dull, T.J.; Chen, EWO J. 6, 3341-3351, 1987
A:Title: Human proto-oncogene c-kit: a new cell surface receptor tyrosine kinase for an A:Reference number: S01426; MUID:88111521; PMID:2448137
A:Accession: S01426
A:Molecule type: mRNA
A:Residues: 1-976 <VAR>
A:Cross-references: GB:X06182; NID:934084; PIDN:CAA29548.1; PID:934085
R:Hu, W.X.; Cornu, F.; Andre, C.; Galibert, F.
Chinese Biochem. J. 7, 618-629, 1991
A:Title: Nucleotide sequence of two neighbouring fragments of human c-kit proto-oncogene A:Reference number: PC1015
A:Accession: PC1015
A:Molecule type: DNA
A:Residues: 412-713 <HOW>
A:Note: article in Chinese with English abstract
R:Spritz, R.A.; Giebel, L.B.; Holmes, S.A.
Am. J. Hum. Genet. 50, 261-269, 1992
A:Title: Dominant negative and loss of function mutations of the c-kit (mast/stem cell 9 A:Reference number: A41815; MUID:92133600; PMID:1370874
A:Accession: A41815
A:Molecule type: DNA
A:Residues: 579-583, 'L', 585-589 <SPR>
A:Cross-references: GB:S78839; NID:9244084; PIDN:AAB21234.1; PID:9244085
A:Note: sequence extracted from NCBI backbone (NCBIN:78839, NCBI:78842)
A:Note: disease-related mutant from patient with piebaldism
A:Accession: B41815
A:Molecule type: DNA
A:Residues: 637-641, 'SPELPW' <SP2>
A:Cross-references: GB:S78843; NID:9244086; PIDN:AAB21235.1; PID:9244087
A:Note: sequence extracted from NCBI backbone (NCBIN:78843, NCBI:78844)
A:Note: disease-related mutant from patient with piebaldism
A:Accession: C41815
A:Molecule type: DNA
A:Residues: 556-560, 'GGDKWK' <SP3>
A:Cross-references: GB:S78845; NID:9244088; PIDN:AAB21236.1; PID:9244089
A:Note: sequence extracted from NCBI backbone (NCBIN:78845, NCBI:78846)
A:Note: disease-related mutant from patient with piebaldism
R:Giebel, L.B.; Strunk, K.M.; Holmes, S.A.; Spritz, R.A.
Oncogene 7, 2207-2217, 1992
A:Title: Organization and nucleotide sequence of the human KIT (mast/stem cell growth fa A:Reference number: I37948; MUID:93064697; PMID:1279499
A:Accession: I37948
A>Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-976 <RES>
A:Cross-references: EMBL:X69301; NID:934089; PIDN:CAA49159.1; PID:9825686
A:Note: an alternative splice form omitting residues 510-513 is described
R:Yamamoto, K.; Tojo, A.; Aoki, N.; Shibuya, M.
Jpn. J. Cancer Res. 84, 1136-1144, 1993
A:Title: Characterization of the promoter region of the human c-kit proto-oncogene. A:Reference number: I56954; MUID:94103107; PMID:7506248
A:Accession: I56954
A>Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-22 <RE2>
A:Cross-references: GB:S67773; NID:9459358; PIDN:AAB29529.1; PID:9459359
R:Spritz, R.A.; Holmes, S.A.; Berg, S.Z.; Nordlund, J.J.; Fukui, K.
Hum. Mol. Genet. 2, 1499-1500, 1993
A:Title: A recurrent deletion in the KIT (mast/stem cell growth factor receptor) proto-c A:Reference number: I54336; MUID:94061059; PMID:7694728
A:Accession: I54336
A>Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 242-250 <RE3>
A:Cross-references: GB:S67686; NID:9460545; PIDN:AAD13996.1; PID:94261696
C:Genetics:
A:Gene: GDB:KIT
A:Cross-references: GDB:120117; OMIM:164920
A:Map position: 4q12-4q12

A;Introns: 23/1; 113/1; 207/1; 252/3; 309/1; 372/2; 411/1; 449/2; 514/1; 549/3; 592/1; 6
A;Note: defects in this gene may result in piebaldism
C;Function:
C;Description: catalyzes the phosphorylation of a peptidyl tyrosine residue by ATP
C;Superfamily: macrophage colony-stimulating factor 1 receptor; immunoglobulin homology;
C;Keywords: alternative splicing; ATP; autophosphorylation; glycoprotein; kinase-related
rotein; tyrosine-specific protein kinase
F;1-976/Product: protein-tyrosine kinase kit precursor, long form #status predicted <MAT
F;1-509,514-976/Product: protein-tyrosine kinase kit precursor, short form #status predi
F;1-22/Domain: signal sequence #status predicted <SIG>
F;23-976/Product: protein-tyrosine kinase kit #status predicted <MAT>
F;23-520/Domain: extracellular #status predicted <EXT>
F;51-99/Domain: immunoglobulin homology <IMM1>
F;129-188/Domain: immunoglobulin homology <IMM2>
F;236-282/Domain: immunoglobulin homology <IMM3>
F;328-394/Domain: immunoglobulin homology <IMM4>
F;423-493/Domain: immunoglobulin homology <IMM5>
F;521-543/Domain: transmembrane #status predicted <TMM>
F;544-976/Domain: intracellular #status predicted <INT>
F;587-931/Domain: protein kinase homology <KIN>
F;595-603/Region: protein kinase ATP-binding motif
F;58-97,136-186,233-290,428-491/Disulfide bonds: #status predicted
F;130,145,283,293,300,320,367,463,486/Binding site: carbohydrate (Asn) (covalent) #s
F;623,640,792/Active site: Lys, Glu, Asp #status predicted
F;797,810/Binding site: magnesium (Asn, Asp) #status predicted

Query Match 22.7%; Score 1196; DB 1; Length 976;
Best Local Similarity 31.4%; Pred. No. 4.7e-54;
Matches 312; Conservative 172; Mismatches 361; Indels 148; Gaps 32;

Qy 36 CVLISHENNGSSAGKPSYRVRGSPEDLOCTPRQSGTVYAAATVBAESGSLTQVQ 95
Db 12 CVLILLRVQTSSGSPS-----VSPGSPSPSPHSGSLDIIVRGD-EIRLL 57

Qy 96 LATPGLDLSCLVFKHSSGCGPHFDLQNGIVSMALNVTTQAGEYLLHIOQSERANYTV 155
Db 58 CTDFGFKV--WTFE---ILDETENKQEWITEKA---EATNTGKYCTNKHGSLNSIY 108

Qy 156 LFTVNVDR-TQYLVRRYFRKMNQDALLCI-----SEG-----VPE 192
Db 109 VF---VRDPKALFLVDRSLYKEDNDTLVRCPDTPDEVNTYSLKCGQKPLPKDLRFIPD 165

Qy 193 PTVEWLVCSHRE-----SCKEEGPAVVRKEKVLHELFGTDIRCCARNALGRECTK 244
Db 166 PKAGIMIKSVKRAVHRLCHLCSVDQEG-----KSVLSEKPFILKVRPAFK-----A 210

Qy 245 LFTIDLNOAPOSTLPQLFLKVGEPILWIRCKAIHVNHGFLGTWLEEDK--ALEE----- 295
Db 211 VPVVSVSKA-----SYLLREGEFTVCTIKDVSSVYTWKRENSQTKLOEKYNSMHH 264

Qy 296 GSYFEMSTYTNRMIRILLAPVSVSGRNDGTYYTCSSKHPSQALVTILE---KGFN 352
Db 265 GDF---NYERQATLT-----ISSRVNDSGVFMCYANNTFGSANVTTLLEVVDKGFN 314

Qy 353 A-TSSQBEYEDPVKFCFSVRFKAYPR-IRCTWIFSQASP--PCEORGLDGYISKFC 408
Db 315 IFPMINTVFDNGENVDLIVEYEAFFPEHQOVIYMNRTFDKWEDEYPKSESNIRIV 374

Qy 409 DHKN-----KPGEYIFEAENDDAQFTKMTLNIRKPKQVLA--NAGASQASCSGDPY 459
Db 375 SELHLTRLKTEGGTYTFLVNSDVAALAFVYVNTPEILTYDLRVNLMQCVAAQFP 434

Qy 460 LPSWTWKCSKSPNCTEEI-PEGVNNKANKRVFGQVSSSTLNSBAGKLLVKCCAY 518
Db 435 EPTIDWYPCPTEQRCSASVLPVDVQTLNSSGPPGKLVQSSIDSSAFKHNGTVECKAY 494

Qy 519 NSMGTSCETIFLNSGPPFPFTQDN-----ISFYATIGLCPLFPVILVILIC 564
Db 495 NDVGKT--SAYFN-----FAFGNNKEQIHPHTLFTPLLIGFVIVAGM---MCIITWILT 544

Qy 565 HXYKQKPYESOLOMI-OVTGPDNDYFVDPDRYEDLKWFEFFRENLEFGRVLGSGAFG 623
Db 545 KYLOKPMYEVQWKVVEING---NNYVIIDTQLPYDHKWEFFRNRLSFGTLGAGAFG 601

Qy 624 RVNATAYGISTGTVSIQVAVKMLKEKADSCKEALMSELKMTHLGHNDINVLIGACT 683
Db 602 KVEATAYGLIKSDAAMTVAVKMLKPSAHLTEREALMSELKVLVYLGNHNVNVLIGACT 661

Qy 694 LSGPVYLIPCYCGDGLNLYLRSKREKPHRTWTEIFKHNHFSYPTFQAHSSNMPGSR 743
Db 662 IGGPTLVITEYCCYDGLNLFNLRKRDSTFCSQEDHAEALYKLLHSSKSSCSDSTNEY 721

Qy 744 VOLHPPDLQD---SGFNGNSHSEDEIEYENQKRLAESEEDLNVLTPEDLICFAYOVAK 800
Db 722 MDMPGVSIVVPTKADKRSVRIGSYIERDVTTPALMEDDELALDL---EDLLSFSYQVAK 778

Qy 801 GMEFLFKSVHRDLAARNVLVTHGKVVKICDPLGLARDILSDSSVYVRGNALPVKWMA 860
Db 779 GWAFILAKNCIHRDLAARNVILTHGRITKICDPLGLARDIKDNTVYVKGNAALPVKWMA 838

Qy 861 ESLFEGYITIKSDVMSYGILLWEIFSLGNVPYGPVDPANFYKLIQSQFMQEPFYATEG 920
Db 839 ESIFNCVYTPESDVMYSYGIFLWELFSLGSSPYGMPVDSKFYKMIKEGFRMLSPHAPAE 898

Qy 921 IYFVMSQWAFDSRKRPSPFNLTSLGCOLABA 953
Db 899 MYDINKTCWDADPLKRPFTFKQIVQLIEKQISES 931

RESULT 14
T30816
macrophage colony-stimulating factor receptor - Japanese pufferfish
C;Species: Fugu rubripes (Japanese pufferfish)
C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 02-Jun-2000
C;Accession: T30816
R;How, G.F.; Venkatesh, B.; Brenner, S.
Genome Res. 6, 1185-1191, 1996
A;Title: Conserved linkage between the pufferfish (Fugu rubripes) and human genes for pla
A;Reference number: Z20882; MUID:97129405; PMID:8973913
A;Accession: T30816
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-975 <HW>
A;Cross-references: EMBL:U63926; NID:G1752706; PID:G1752708; PIDN:AAC60063.1
C;Superfamily: macrophage colony-stimulating factor 1 receptor; immunoglobulin homology;

Query Match 22.6%; Score 1188.5; DB 2; Length 975;
Best Local Similarity 31.1%; Pred. No. 1.1e-53;
Matches 318; Conservative 173; Mismatches 336; Indels 195; Gaps 31;

Qy 74 GTVYEAAVTE-----VAESGSLTQ-----VOLATPGDLSCLW---VFKHSSLCG 115
Db 10 GIMASASSVEWRHPVWVFNKSVQSSVVVVGTSLELKCQGGDGPVNWQTRLPKHKRYMS 69

Qy 116 QPHFDLQNRGIVSMALNVTTQAGEY--LLHIOQSERANYTVLFTVNVDR-----T 164
Db 70 RSPGNLR-----TIRVARPTAEFTGTYKCFYSAWAQHRLTSSVHVYVKDPNRFVWTSST 124

Qy 165 QLYVLRFRYFRKMNQDALLCISEGVPPTVEWLVLCSSHRECKEEGPAVVRKEKVLHE 224
Db 125 SLRVVRK-----EGEDYLLPCLLTDPDPEATDLGLRMDNGTTPVPEMNTVYVHRGLIRS 178

Qy 225 L---FGTDIRCCARNALGRECTKLTIDLNOAPOSTLPQLFLK-----VCEPLWIRCK 274
Db 179 LQPSFNADVCTAKVKGEKTSKTSFINVIQKLRFP-PYVLEMDYVYRIVGEEIQIRCM 237

Qy 275 AIHVNHGFLGTWLEEDKA---LEEGSFEMSTYTNRMIRILLAFVSVSGRNDGTYYTC 331
Db 238 THNPNFNVNTWNTYTKSRVIEE---RVSSSGENRLDIQSILT-ISAVDLADTGNISC 292

Qy 332 SSSKHP-----SQSALVTILEKGF-----NATSSQBEYEDPVKFCFSVRFKAYPRI 380
Db 293 IGTNEAGVNSNTVLLVVEKPYIRLWLPQLIPKLASQGLSVENEGEDLELGVVVEAYPQI 352

Qy 381 RCTWIFSQASPCQORGLDGYISKFCDHK-----NK 413

Db 353 -----TDHRWHPTSPSTMQEHIYHARLQKRMNAQE 385
Qy 414 PGEYFYAENDDAQTKMTLIRKPK-QVLNANASQAQSSDGGVPLPSMTWKCSKDS 472
Db 386 QGYTYFAKSNLANGSIHFVKMYQKPAVVRWENITLTCTSGYPAPAIWIYQCSGIR 445
Qy 473 PNCT-----EETPEGVWKKANRKFVQGWSSSTLNMSSBAGKGLLVKCCAYNSMGTS 524
Db 446 PTCNGNNTGLPKQNHPOAL-TVEVQREYGAVEVESVFTVGLSNHRMTVECVAFNLVGV 504
Qy 525 CETFLNSPGPPFFODNISFATTGLC--LPFIVLVILCHKKQKQFRIESQLOMIQV 582
Db 505 SDTFTVE-----VSDLFTSTLCAAGVLAIFLLLVFLYKQKQKQFRIKRIIEA 557
Qy 583 TGPLDNEYFYVDFRDEYDLKWEPPRENLKFGKVLGSGAFGVMNATAYGISK-TGVSIQ 641
Db 558 RE--GNNTFTDPTOLPYNEKWEPRDKLKLKVLGAGAFKVVATAFGLGEDKXNLTIR 615
Qy 642 VAVKMLKEKADSCKEALMSELKMMTHLGHHDNIIVNLGACTSLGPFVLIPIFYCCYGDLL 701
Db 616 VAVKMLKANAHSDEREALMSELKILSHLGHQNIIVNLGACTYGGFVLVITEYCSLGDLL 675
Qy 702 NYLRSKREKPHRTWTEIPKEHNFSSYPTQASHN-----SSMPGSE 743
Db 676 NFLROKAETFEVNLVMI-----PEIMNSNDYKNIQKMYIRSDSGISSTSSY 726
Qy 744 VOLHPLDQLSGFNGNSIHSEDEIEVENOKRLAEBEEDNLVLTPEDLICFAYQVAKGME 803
Db 727 LEMRP-----SQSHIISAGRSKSLCEBNGD--WFLDIDDLRLFSLOVAQGLD 771
Qy 804 FLEFKSCVHRDLAARNVLVTHGKVKI CDPLGLARDILSDSSYVVRGNARLPVKWMAPESL 863
Db 772 FLASRNCIHRDVAARNVLLTDKRAKI CDPLGLARDIMDSNVVVGNAARLPVKWMAPESI 831
Qy 864 PEGYITISDWSYGIILLWEIPSLGVPYGPVDPANFYKLIQSFGKMEQPYATGFIYF 923
Db 832 FDCVYTVQSDWSYGIILLWEIPSLGKSPYPSVAVDSRFYMKRGYQMSQDPFALPEIYM 891
Qy 924 VQWSCAFDSRKPFPNLTSLFLGCOLA--EAEEACIRTSIHLPKQAAPOQRGLRAQSP 981
Db 892 IMKWCWNLPTERTPTFSMISQWNLRLGGODEKELIIRNVO-PEQVAE-----GEACDEP 946
Qy 982 QR 983
Db 947 KR 948

RESULT 15
JN0677
protein-tyrosine kinase (EC 2.7.1.112), receptor type kit precursor - chicken
N:Alternate names: tyrosine kinase receptor kit
C:Species: Gallus gallus (chicken)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 16-Jun-2000
C:Accession: JN0677
Gene 128, E.; Okamura, H.; Chikamura, T.; Kanai, Y.; Watanabe, M.; Naito, M.; Sakurai, M.
A:Title: Cloning and expression of the chicken c-kit proto-oncogene.
A:Reference number: JN0677; PMID:93292995; PMID:7685729
A:Accession: JN0677
A:Molecule type: mRNA
A:Residues: 1-960 <SAS>
A:Cross-references: DBJ:D13225; NID:G303532; PID:BAA02506.1; PID:G303533
A:Experimental source: brain
C:Superfamily: macrophage colony-stimulating factor 1 receptor; immunoglobulin homology;
rotein kinase
F:1-24/Domain: signal sequence #status predicted <SIG>
F:25-960/Product: tyrosine kinase receptor #status predicted <MAT>
F:314-380/Domain: immunoglobulin homology <IMM>
F:573-916/Domain: protein kinase homology <KIN>
F:581-589/Region: protein kinase ATP-binding motif
F:76,135,149,269,286,308,318,338,356,453,469/Binding site: carbohydrate (Asn) (covalent)

Query Match 22.4%; Score 1181; DB 1; Length 960;
Best Local Similarity 31.8%; Pred. No. 2.7e-53;
Matches 310; Conservative 159; Mismatches 356; Indels 150; Gaps 30;
Qy 45 GSSACKPSYRMVRGSPEDLOCTPRQSEGTVEAATVEVAESGSITLQVQLATGDLCL 104
Db 24 GSVPHSESLVNVKGEELKCN-----EEGPTVWNFQNSDPSAKTR 65
Qy 105 L-----WFKHSSLCQPHFDLQNRGIVSMALINVTETQAGEYLLHIOSEARANYTVLFTV 159
Db 66 ISNEKEWTKNATRIDIGRYECKSG-----SLVN-----SFIV 99
Qy 160 NVRDQLVLRARPYRKMENODALCISEGVPEPTVFWLCSHRECKEBGAVVRKEE 219
Db 100 FVKDPNVLFLVDSLITYGKEDSDILLVCLTPDP-VLNFETLRKCDGKPLPKNMTFIPNPK 158
Qy 220 ----KVLHELFGTDIRCCAR-NALGRECTKLFTIDLNOAP-QSTLPOL-----FLKV 266
Db 159 GIIHKVQSRPKGYCQCLAKHNGVEKISEHIP---LNVRPVHKALPVLITLSKSYELLKEG 215
Qy 267 EPLWIRKCAIHWNGHFGLTWELEDKALBEGSYFEMSTYSTNRMTIRIILAFVSSVGRNDT 326
Db 216 BEFVETCIITDVS SVKASWISYKSAIVTSKRNLDGYERK----LTNIRSVGVNDS 271
Qy 327 GYTCSSSKHP--SOSALVTI--LEKGFINATSOE-EYEIDPYKFCFVSFRKAYPRIR 381
Db 272 GEFTC-QAENPFKTNATVTLKALAKGFVRLPATWNTTIDINAGQGNLTVYEAYPK-- 328
Qy 382 CTWIFSOASPFCEQGLDGYGISKFDHKNK-----PGEYIF 419
Db 329 -----PKEEYVMYMNETLQNSDHYKFKTVGNNSYTSSELHTRLKGTGEGIYTF 378
Qy 420 YAENDDAQTKMTLIRKPKQVLNANASQ--ASCSSDGGVPLPSMTWKCDKSPNCTE 477
Db 379 FVNSDASSVTFNYYKTKPEILLTDLMLGNDILQCVATGFPAPTIYWFCEGTGECRLD 438
Qy 478 E---IPEGVWKKANRKY--FGQWSSSTLNMSEAGKGLLVKCCAYNSMGTSCEITFLNS 532
Db 439 SPTISPMKVSYNTSSVPSFERILVESTVNAS-MFKSTGTICCEASSNGDK-SSVFFNF 496
Qy 533 PGPPFFIODNISFYA-----TGLCLPFIIVLVILCHKKQKQFRIESQLOMI-QVTG 584
Db 497 A-----IKEQIRTHLTFTPLLIAGVAAAGLCIIVMILVYIYLOPKYEVQWKVVEEING 551
Qy 585 PLDNEYFYVDFRDEYDLKWEPPRENLKFGKVLGSGAFGVMNATAYGISKTVSIQAV 644
Db 552 ---NNYVIDPTQLPYDHKWEPPRNLISFGKTLGAGAFKVVATAYGLFSDAAMTVAV 608
Qy 645 KMLKEKADSCKEALMSELKMMTHLGHHDNIIVNLGACTSLGPFVLIPIFYCCYGDLLNYL 704
Db 609 KMLKPSAHLTREALMSELKVLVSLGNHINIVNLGACTIGGPTLIVITEYCCYGDLLNLF 668
Qy 705 RSKREKP-----HRTWTE--IPKEHNFSSYPTQOA-HSNSSMPGSRVQLHPPLDQLSGFN 757
Db 669 RRKRDSTFCPCHHEEAHAAVYENLHQAEPTADAVNEMDMKPGVSYAVPPKADKKRPVK 728
Qy 758 GNSIHSEDEIEVENOKRLAEBEEDNLVLTPEDLICFAYQVAKGMEFLEFKSCVHRDLAA 817
Db 729 SGS-----YTQDDVTLSMLEDELALDELDLVSFQYVAKGMSFLASKNCHIRDLAA 780
Qy 818 RNVLVTHGKVKI CDPLGLARDILSDSSYVVRGNARLPVKWMAPESLFEGIYTIKSDVMSY 877
Db 781 RNILLTHGRITKI CDPLGLARDIRNDSNVVVGNAARLPVKWMAPEISFNCVYTFESDVMSY 840
Qy 878 GILWEIPSLGVPYGPVDPANFYKLIQSFGKMEQPYATGFIYFVWQSCWAFDSRKP 937
Db 841 GILLWEIPSLGSSPYGPMVDSKPYKMKIEGVRMPSPECSPPMYDINKSCWDADPLQRP 900
Qy 938 SFPNLTSLGCOLAE 952
Db 901 TFKQIVQLIEQOLSD 915

Search completed: August 26, 2003, 07:29:00
Job time : 36.9834 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2003, 07:14:08 ; Search time 18.9904 Seconds
(without alignments)
2456.527 Million cell updates/sec

Title: US-09-919-408A-2

Perfect score: 5264

Sequence: 1 MRALAQRGDRRLLLWLSV.....RGLRAQSPQOVKIHRS 992

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5264	100.0	992	1 FLT3 MOUSE	Q00342 mus musculus
2	4429.5	84.1	993	1 FLT3 HUMAN	P36888 homo sapien
3	1258.5	23.9	980	1 KFN5_FELCA	P13369 felis silve
4	1249.5	23.7	978	1 KFN5_FSVMD	P00545 feline sarc
5	1223	23.2	975	1 KIT MOUSE	P05532 mus musculus
6	1222.5	23.2	972	1 KIT CAPHI	Q28317 capra hircu
7	1221.5	23.2	978	1 KFN5 HUMAN	P07333 homo sapien
8	1219	23.2	977	1 KFN5 MOUSE	P09581 mus musculus
9	1215	23.1	977	1 KIT BOVIN	P43481 bos taurus
10	1210	23.0	978	1 KFN5 RAT	Q00495 rattus norv
11	1207	22.9	975	1 KIT CANFA	Q97799 canis famil
12	1196	22.7	976	1 KIT HUMAN	P10721 homo sapien
13	1181	22.4	960	1 KIT CHICK	Q08156 gallus gall
14	1178.5	22.4	978	1 KIT FELCA	Q28889 felis silve
15	1157.5	22.0	1088	1 PGDS RAT	P20786 rattus norv
16	1152.5	21.9	1089	1 PGDS HUMAN	P16234 homo sapien
17	1142.5	21.7	1089	1 PGDS MOUSE	P26618 mus musculus
18	1132	21.5	1087	1 PGDS XENLA	P26619 xenopus lae
19	1098	20.9	1098	1 PGDR MOUSE	P05622 mus musculus
20	1078	20.5	1106	1 PGDR HUMAN	P09619 homo sapien
21	998.5	19.0	1338	1 VGR1 HUMAN	P17948 h vascular
22	984.5	18.7	1336	1 VGR1 RAT	P53767 rattus norv
23	983.5	18.7	1333	1 VGR1 MOUSE	P35969 mus musculus
24	967.5	18.4	370	1 KIT FSVHZ	P04048 feline sarc
25	965	18.3	1348	1 VGR2 COTVA	P52583 coturnix co
26	961.5	18.3	1356	1 VGR2 HUMAN	P35968 homo sapien
27	954.5	18.1	1298	1 VGR3 HUMAN	P35916 homo sapien
28	952	18.1	1363	1 VGR3 MOUSE	P35917 mus musculus
29	934	17.7	1343	1 VGR2 RAT	O08775 rattus norv
30	926.5	17.6	1367	1 VGR2 MOUSE	P35918 mus musculus
31	818	15.5	813	1 FGR2 XENLA	Q03364 xenopus lae
32	810.5	15.4	823	1 CEK3_CHICK	P18461 gallus gall
33	791	15.0	821	1 FGR2_HUMAN	P21802 homo sapien

34	789	15.0	806	1 CEK2_CHICK	P18460 gallus gall
35	780	14.8	821	1 FGR2 MOUSE	P21803 mus musculus
36	766.5	14.6	819	1 FGR1_CHICK	P21804 gallus gall
37	761.5	14.5	654	1 BFR2_HUMAN	Q01742 homo sapien
38	758.5	14.4	812	1 FGR1_XENLA	P22182 xenopus lae
39	754.5	14.3	822	1 FGR1_HUMAN	P11362 homo sapien
40	751.5	14.3	806	1 FGR3_HUMAN	P22607 homo sapien
41	751.5	14.3	822	1 FGR1_MOUSE	P16092 mus musculus
42	748.5	14.2	822	1 FGR1_RAT	Q04589 rattus norv
43	732	13.9	801	1 FGR3_MOUSE	Q61851 mus musculus
44	698	13.3	802	1 FGR4_HUMAN	P22455 homo sapien
45	697	13.2	808	1 FGR4_MOUSE	Q03142 mus musculus

ALIGNMENTS

RESULT 1
FLT3_MOUSE
ID FLT3_MOUSE STANDARD; PRT; 992 AA.
AC Q00342;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE FL cytokine receptor precursor (EC 2.7.1.112) (Tyrosine-protein kinase receptor flk-2) (Fetal liver kinase 2) (Tyrosine-protein kinase FLT3).
GN FLT3 OR FLT-3 OR FLK-2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=91292518; PubMed=1648448;
RA Matthews W., Jordan C.T., Wiegand G.W., Pardoll D., Lemischka I.R.;
RT "A receptor tyrosine kinase specific to hematopoietic stem and progenitor cell-enriched populations.";
RL Cell 65:1143-1152(1991).
[2]
RP SEQUENCE FROM N.A.
RX MEDLINE=92019834; PubMed=1656368;
RA Rosnet O., Marchetto S., Delapeyriere O., Birnbaum D.;
RT "Murine Flt3, a gene encoding a novel tyrosine kinase receptor of the PDGFR/CSFIR family.";
RL Oncogene 6:1641-1650(1991).
[3]
RP CHARACTERIZATION.
RX MEDLINE=93205405; PubMed=8384358;
RA Maroc N., Kottapel R., Rosnet O., Marchetto S., Lavezzi C., Mannoni P., Birnbaum D., Dubreuil P.;
RT "Biochemical characterization and analysis of the transforming potential of the FLT3/FLK2 receptor tyrosine kinase.";
RL Oncogene 8:909-918(1993).
CC -!- FUNCTION: RECEPTOR FOR THE FL CYTOKINE. HAS A TYROSINE-PROTEIN KINASE ACTIVITY.
CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein tyrosine phosphate.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- TISSUE SPECIFICITY: HEMATOPOIETIC STEM AND PROGENITOR CELL-ENRICHED POPULATIONS. FOUND IN BRAIN, PLACENTA AND TESTIS.
CC -!- SIMILARITY: BELONGS TO THE CSF-1/PDGFR RECEPTOR FAMILY OF TYROSINE-PROTEIN KINASES.
CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL; M64689; AAA37634.1; --
 DR EMBL; X59398; CAA42041.1; --
 DR PIR; A39931; A39931.
 DR PIR; S18827; S18827.
 DR HSP; S11362; 1FGK.
 DR MGD; MGI:95559; Flt3.
 DR InterPro; IPR007110; Ig-like.
 DR InterPro; IPR003599; Ig.
 DR InterPro; IPR003006; Ig.MHC.
 DR InterPro; IPR000719; Prot kinase.
 DR InterPro; IPR001824; RTKinaseII.
 DR InterPro; IPR001245; Tyr_kinase.
 DR Pfam; PF00047; Ig; 1.
 DR Pfam; PF00069; pkinase; 1.
 DR ProDom; PD000001; Prot kinase; 2.
 DR SMART; SM00219; TyKc; 1.
 DR PROSITE; PS50835; IG-LIKE; 1.
 DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
 DR PROSITE; PS00111; PROTEIN KINASE DOM; 1.
 DR PROSITE; PS00109; PROTEIN KINASE TYR; 1.
 DR PROSITE; PS00240; RECEPTOR TYR KIN III; 1.
 KW Signal; Transferase; Tyrosine-protein kinase; Receptor; Transmembrane;
 KW Glycoprotein; Phosphorylation; ATP-binding; Immunoglobulin domain.
 FT SIGNAL 1 27 POTENTIAL.
 FT CHAIN 28 992 FL CYTOKINE RECEPTOR.
 FT DOMAIN 28 544 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 545 564 POTENTIAL.
 FT DOMAIN 565 992 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 254 344 IG-LIKE C2-TYPE.
 FT DOMAIN 611 946 PROTEIN KINASE.
 FT NP_BIND 617 625 ATP (BY SIMILARITY).
 FT BINDING 645 645 ATP (BY SIMILARITY).
 FT ACT_SITE 814 814 BY SIMILARITY.
 FT CARBOHYD 44 44 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 133 133 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 152 152 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 307 307 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 324 324 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 352 352 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 445 445 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 474 474 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 503 503 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 542 542 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CONFLICT 150 150 R -> A (IN REF. 2).
 FT CONFLICT 242 242 C -> S (IN REF. 2).
 FT CONFLICT 726 726 S -> F (IN REF. 2).
 FT CONFLICT 957 979 CRTSLHLPKQAPQORGGRLAQ -> MYQNMGGNVPEHPS
 FT CONFLICT 983 983 IYONRRPLSRAGSEPP (IN REF. 2).
 FT CONFLICT 983 983 R -> A (IN REF. 2).
 SQ SEQUENCE 992 AA; 112639 MW; 407A087853372100 CRC64;

Query Match 100.0%; Score 5264; DB 1; Length 992;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 992; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRALQSRDRRLLLVLSVMILETTVNDLPVVKCVLISHENNGSSAGKPSYRMVRS 60
 DB 1 MRALQSRDRRLLLVLSVMILETTVNDLPVVKCVLISHENNGSSAGKPSYRMVRS 60
 QY 61 PEDLOCTPRQSEGTVEAATVEAAGSITLQVLAATPDGLSCLVFKHSLGCGPHFD 120
 DB 61 PEDLOCTPRQSEGTVEAATVEAAGSITLQVLAATPDGLSCLVFKHSLGCGPHFD 120
 QY 121 LQNRGIVSMAILNVETQAGEYLLHHTQSERANYTLFTVNVRTQLYVLRPFRQWENQ 180
 DB 121 LQNRGIVSMAILNVETQAGEYLLHHTQSERANYTLFTVNVRTQLYVLRPFRQWENQ 180
 QY 181 DALLCISEGVPETVEMVLCSSHRECKEKGPAVVRKEEKLHLEFGTDIRCCARNALGR 240
 DB 181 DALLCISEGVPETVEMVLCSSHRECKEKGPAVVRKEEKLHLEFGTDIRCCARNALGR 240
 QY 241 ECTKLTIDLNOAPOSTLPQLFLKVGEPWLIRCKAIHNVHFGTLWELEDKALEEGSYFE 300

DB 241 ECTKLTIDLNOAPOSTLPQLFLKVGEPWLIRCKAIHNVHFGTLWELEDKALEEGSYFE 300
 QY 301 MSTYSTNRTMIRILLAFVSSVGRNDTGYTSCSSKHPQSALVTILEKGFINATSOEEY 360
 DB 301 MSTYSTNRTMIRILLAFVSSVGRNDTGYTSCSSKHPQSALVTILEKGFINATSOEEY 360
 QY 361 EIDPYEKFCFSVRFKAYPRIRCTWIFSQASFFCEQRLGLEDGYISISKFCDHKNKPGYIFY 420
 DB 361 EIDPYEKFCFSVRFKAYPRIRCTWIFSQASFFCEQRLGLEDGYISISKFCDHKNKPGYIFY 420
 QY 421 AENDDAQFTKMTLIRKPKQVLANASASQASCSDDGYPLPSWTWKCKSDKSNCTEEIP 480
 DB 421 AENDDAQFTKMTLIRKPKQVLANASASQASCSDDGYPLPSWTWKCKSDKSNCTEEIP 480
 QY 481 EGVNKKANRKFVGOWVSSSTLNMSEAGKLVKCCAYNSMGTSCTIFLNSGPPFPFIQ 540
 DB 481 EGVNKKANRKFVGOWVSSSTLNMSEAGKLVKCCAYNSMGTSCTIFLNSGPPFPFIQ 540
 QY 541 DNISFYATIGLCLPFIVLVLVILCHIKYKQPRYESQLOMIQVTGPLDNEYFYVDFRDY 600
 DB 541 DNISFYATIGLCLPFIVLVLVILCHIKYKQPRYESQLOMIQVTGPLDNEYFYVDFRDY 600
 QY 601 DLKWFPPRENLEFGKVLGSGAFGRVNMATAYGISKTGVSIOAVKMLKEKADSCKEALM 660
 DB 601 DLKWFPPRENLEFGKVLGSGAFGRVNMATAYGISKTGVSIOAVKMLKEKADSCKEALM 660
 QY 661 SELKWMTHLGHNDINVLGACTLSGPVYLIFEYCCYGDLLNLYLRKREKFRHTWTWEIFK 720
 DB 661 SELKWMTHLGHNDINVLGACTLSGPVYLIFEYCCYGDLLNLYLRKREKFRHTWTWEIFK 720
 QY 721 ENHFSYPTFOAHSNSMPGSRVQLHPDLQSLGFNGNSIHSEDEIYEYENQKRLAESEE 780
 DB 721 ENHFSYPTFOAHSNSMPGSRVQLHPDLQSLGFNGNSIHSEDEIYEYENQKRLAESEE 780
 QY 781 EDNLVLTTFEDLLCFAYQVAKMEFLFKSCVHRDLAARNVLTGHGVKVICDFGLARDIL 840
 DB 781 EDNLVLTTFEDLLCFAYQVAKMEFLFKSCVHRDLAARNVLTGHGVKVICDFGLARDIL 840
 QY 841 SDSSYVVRGNARLPVKWMAPESLFEGIYTKSDVMSYGILLWEIFSLGVNYPYGPVPDAN 900
 DB 841 SDSSYVVRGNARLPVKWMAPESLFEGIYTKSDVMSYGILLWEIFSLGVNYPYGPVPDAN 900
 QY 901 FYKLQSGFKMEQPFYVATYEGIFVMSQWAFDSRKRKPSFNLTSFLGCOLAEAEACIRT 960
 DB 901 FYKLQSGFKMEQPFYVATYEGIFVMSQWAFDSRKRKPSFNLTSFLGCOLAEAEACIRT 960
 QY 961 SIHLPKQAPQORGGRLAQSPOQVKIHRERS 992
 DB 961 SIHLPKQAPQORGGRLAQSPOQVKIHRERS 992

RESULT 2
 FLT3_HUMAN STANDARD; PRT; 993 AA.
 ID AC P36888; Q13414;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE FL cytokine receptor precursor (EC 2.7.1.112) (Tyrosine-protein kinase
 GN receptor FLT3) (Stem cell tyrosine kinase 1) (STK-1) (CD135 antigen).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94119906; PubMed=7507245;
 RA Small D., Levenstein M., Kim E., Garow C., Amin S., Rockwell P.,
 RA Witte L., Burrow C., Ratajczak M.Z., Gewirtz A.M., Civin C.I.;
 RT "STK-1, the human homolog of Flk-2/Flt-3, is selectively expressed in
 CD34+ human bone marrow cells and is involved in the proliferation of


```

Db 660 SELKMTQLGSHENIVNLGACTLSGPIYLIFEYCCYGLNLYRSKREKPHRTWTETFK 719
Qy 721 EHNFSYPTFOAHNSMPSGSEVQLHPDLQSLGFGNGSIHSESDIEYENOKRLAEERE 780
Db 720 EHNFSYPTFOAHNSMPSGSEVQLHPDQSLGFGNGSFHSESDIEYENOKRL--EE 777
Qy 781 EDNLVLTEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVVKICDFGLARDIL 840
Db 778 EDNLVLTEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVVKICDFGLARDIM 837
Qy 841 SSSSVTVRGNARLPVKWAPESLFEIGIYTIKSDVMSYGILLWEIISLGNVPGIPVDAN 900
Db 838 SDSNTVVRGNARLPVKWAPESLFEIGIYTIKSDVMSYGILLWEIISLGNVPGIPVDAN 897
Qy 901 FYKLLQSGFKMQPFPVATGEGIVFMQSCWAFDSRKPSPNLTSLFGCOLAABEAC--- 957
Db 908 FYKLLQSGFKMQPFPVATGEGIVFMQSCWAFDSRKPSPNLTSLFGCOLAABEAMQYN 957
Qy 958 ----IRTSIHLPKQAPQORG-GLRAQSPQRQVK 986
Db 958 VDRVSECPHTYQNRFPSPREMDLGLLSPQAQVE 991

RESULT 3
KFMS_FELCA
ID_KFMS_FELCA STANDARD; PRT; 980 AA.
AC P13369;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Macrophage colony stimulating factor I receptor precursor (CSF-1-R)
DE (EC 2.7.1.112) (fms proto-oncogene) (c-fms).
GN CSF1R OR FMS.
OS Felis silvestris catus (Cat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
OX NCBI TaxID=9685;
RN [1]_TaxID=9685;
RP SEQUENCE FROM N.A.
RX MEDLINE=89077553; PubMed=2849512;
RA Woolford J., McAluliffe A., Rohrschneider L.R.;
RT "Activation of the feline c-fms proto-oncogene: multiple alterations
are required to generate a fully transformed phenotype.";
RL Cell 55:965-977(1988).
CC -!- FUNCTION: THIS PROTEIN IS THE RECEPTOR FOR CSF-1, IT IS A PROTEIN
CC -!- TYROSINE-KINASE TRANSMEMBRANE RECEPTOR.
CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-
CC -!- PROTEIN KINASES.
CC -!- SIMILARITY: Contains 5 immunoglobulin-like C2-type domains.
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; J03149; AAA30811.1; -.
CC PIR; A11636; TVCTMD.
CC HSP; P11362; IFKC.
CC InterPro; IPR007110; Ig-like.
CC InterPro; IPR003598; Ig_C2.
CC InterPro; IPR003006; Ig_MHC.
CC InterPro; IPR000719; Prot_kinase.
CC InterPro; IPR001824; RTKinaseII.
CC InterPro; IPR001245; Tyr_pkinase.
CC Pfam; PF00047; Ig; 3.
CC Pfam; PF00069; pkinase; 1.

```

```

DR ProDom; PD000001; Prot_kinase; 2.
DR SMART; SM00408; IGC2; 1.
DR SMART; SM00219; TyrKc; 1.
DR PROSITE; PS00835; IG LIKE; 3.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE; PS00240; RECEPTOR_TYR_KIN_III; 1.
KW Proto-oncogene; Tyrosine-protein kinase; Receptor; Transmembrane;
KW Transferase; Glycoprotein; Phosphorylation; ATP-binding; Signal;
KW Immunoglobulin domain; Repeat.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 980 MACROPHAGE COLONY STIMULATING FACTOR I
FT RECEPTOR.
FT EXTRACELLULAR (POTENTIAL).
FT DOMAIN 20 509 POTENTIAL.
FT TRANSMEM 510 535 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 536 980
FT DOMAIN 21 100 IG-LIKE C2-TYPE 1.
FT DOMAIN 107 197 IG-LIKE C2-TYPE 2.
FT DOMAIN 202 297 IG-LIKE C2-TYPE 3.
FT DOMAIN 299 397 IG-LIKE C2-TYPE 4.
FT DOMAIN 400 499 IG-LIKE C2-TYPE 5.
FT DOMAIN 579 908 PROTEIN KINASE.
FT NP_BIND 585 593 ATP (BY SIMILARITY).
FT BINDING 613 613 ATP (BY SIMILARITY).
FT ACT_SITE 776 776 BY SIMILARITY.
FT DISULFID 42 84 POTENTIAL.
FT DISULFID 127 177 POTENTIAL.
FT DISULFID 224 278 POTENTIAL.
FT DISULFID 417 482 POTENTIAL.
FT MOD_RES 807 807 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT CARBOHYD 45 45 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 73 73 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 94 94 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 153 153 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 275 275 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 286 286 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 302 302 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 335 335 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 410 410 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 477 477 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 490 490 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 980 AA; 108506 MW; 4B5CF6E1B97CF6FF CRC64;

Query Match 23.9%; Score 1258.5; DB 1; Length 980;
Best Local Similarity 32.7%; Pred No. 8.9e-77;
Matches 328; Conservative 168; Mismatches 353; Indels 153; Gaps 32;

Qy 75 TVYEAAATVEVAESGSITLQVQLATPGDLISCL-----W---VFKHSSLGCQPHFDLQNRG 125
Db 13 TAWHAQGVPIQSPGPELVVEPGTTVTLRCVGVSGVEMDGPISPHWNLDLDPSSI---- 68
Qy 126 IVSMALINVTETQAGEYLLHI-----QSERANTVTLFTVNVROTQLVYL---RRPY----- 173
Db 69 ---LTTNNATFQNTGTY--HCTEPGNPGGNATI-----HLVYKDPARPKVLAQE 114
Qy 174 FRKMENQDALL-CISEGPEPTVEVLCSSHRESCKECPAVVRKEEVLHELFGTDD1-- 230
Db 115 VTVLESGDALLPCL---LTDPALE-----AGVSLVRVRGRPVLRQTNVTSFSPWHGTTHK 166
Qy 231 -----RCCARNALGRECTKL-----FTTDLNQAPOSTL--POLFLKV-GEPLWIRCK 274
Db 167 AKFIENHYVQCSAR-VDRGTVTSMGLWKVKOKDISGPATLTLEPAELVRIQGEAAQIVCS 225
Qy 275 AIHVNHGFLTWELDKALEGSEYFEMSTYSTNTRMIRILLAFVSVSGRNDTGYTCSSS 334
Db 226 ASNIDVNFDFVSLRHGDTKL---TISQSDFDHNRVQ-KVLTNLNLDHVSFDAGNYSCTAT 281
Qy 335 K---HPSQSALVTILEKGFINATSSQE-EYEIDPVEKFCFSVRFKAYPIRRC-TWI---- 385
Db 282 NAWGNHASMVPRVVSAYLNLTSEQSLLQEVTVGKVDLQKVAYPGLSEFNTWYLP 341
Qy 386 FSQASFPCEQRGLEDDGYSISKPCD-----HKNKPGEYIFYAENDDAQFTKMTLNIRKKPQ 441

```

```

Db 342 FSDYQDKLDFVIKTYRTYRSTLSLRLKRSSEAGYRSLRNAGGQNALTFELTLRYPE 401
Qy 442 VLANASASQAS---CSSDGYPLPSWTWKCKSDKSPNCTEE---IPEGVWNKANKRVFG 494
Db 402 VRVWTTLINGSDTLLCEASGYPOPSVTWQCRSHTRDCDESAGLVLEDHSHSVLSQVPH 461
Qy 495 QWVSSTLNMSSAGKGLLVKCCAYNSMGTSCTIFLNSPGPPFFQDNISFYATIGLC-- 552
Db 462 EVIVHSLLAIGTLEHNRTYCEAFNSVGNSSQTFPWPISGAHTQPLDELLTFVLVTCMS 521
Qy 553 -LPFTWLIVLCHYKKGFYRYSOLQMTQVGTPLDNEYFYDFRDEYEDLKWEPFRENL 611
Db 522 IMALLLLLLLLLLLYKKQPKQVRWKIIE--SYEGNSTFTDPTQLPYNKWEFPRNLL 579
Qy 612 EFGKVLGAGFGRVNMATAYIGSKTGVSIQAVKMLKEKADSEKALMSLKMTHLGH 671
Db 580 QFGKTLGAGAFGKVEATAFGLGKEDAVLKVAVMKLSKTAHADEKALMSKLSHLQ 639
Qy 672 HDNIYNLLGACTLSGRVYLIFCYCCYGDLLNVLRSKRE-----KF 711
Db 640 HENIYNLLGACTHGGFVLVITYCCYGDLLNLRQAEAMLGPSLSVGQDPAGAGYKNI 699
Qy 712 HRTWTEIFKEHNFSSVPTFOAHNSNMPGSRVQLHPPLDOLSGFNGNSIHSEDELEYEN 771
Db 700 HLEKKYVRKSDFS-----SQGVDTVEMRPVSTSSNDSFSEEDL----- 740
Qy 772 QKRLAEEBEDNLVLTFFDLFCAYQVAKMGFELEPKSCVHRDLAARNVLVTHGKVYIC 831
Db 741 -----GKEDGRPLERLDLHFSSQVQAQMAFLASKNCIHRDVAARNVLLTSGRVAKIG 793
Qy 832 DFLGARDILSDSVYVGNARLPVKMAPESIFEGIYTIKSDVMSYGIILLWEIFSLGVNP 891
Db 794 DFLGARDIMDSNYIVKGNARLPVKMAPESIFDCVYTVQSDVMSYGIILLWEIFSLGLNP 853
Qy 892 YGIPVDANFYKLIQSGFMQEPFYATEGTYFMQSCWAFDSKRSFENLTSFLGCOLA 951
Db 854 YPGILVNSFKVLKVDGYQMAQPAFAPAKPNYSIMQACWALEPRTFRFOOICSL--QK 911
Qy 952 EAEACIRTSIHLKQAAPOQ-----RGLRAQSPQ 984
Db 912 AQEDRRVPNTNLPSSSSSSSSSSSSSRTSGTSGGSSSEPEEE 953

RESULT 4
ID KEMS FSVMND STANDARD; PRT; 978 AA.
AC P00545; Q86597;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Tyrosine-protein kinase transforming protein fms (EC 2.7.1.112).
GN V-FMS.
OS Feline sarcoma virus (strain McDonough).
OC Viruses; Retroviral viruses; Retroviridae; Mammalian type C retroviruses.
OX NCBI_TaxID=11778;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=84119469; PubMed=6582485;
RA Hampe A., Gobet M., Sherr C.J., Galibert F.;
RT "Nucleotide sequence of the feline retroviral oncogene v-fms shows
RT unexpected homology with oncogenes encoding tyrosine-specific protein
RT kinases.";
RL Proc. Natl. Acad. Sci. U.S.A. 81:85-89(1984).
RN [2]
RP REVISIONS, SEQUENCE FROM N.A.
RX MEDLINE=92015516; PubMed=1833563;
RA Smola U., Hennig D., Hadwiger-Fangmeier A., Schuetz B., Pfaff E.,
RA Niemann H., Tamura T.;
RT "Reassessment of the v-fms sequence: threonine phosphorylation of the
RT COOH-terminal domain";
RL J. Virol. 65:6181-6187(1991).
CC -!- FUNCTION: V-FMS IS DERIVED FROM THE RECEPTOR FOR COLONY

```

```

CC STIMULATING FACTOR 1 (CSF-1).
CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC tyrosine phosphate.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- MISCELLANEOUS: THIS PROTEIN IS SYNTHESIZED AS A GAG-FMS
CC POLYPROTEIN.
CC -!- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-
CC PROTEIN KINASES.
CC -!- SIMILARITY: Contains 5 immunoglobulin-like C2-type domains.
CC -----
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; K01643; AAA43045.1; -.
CC EMBL; S59588; AAB20028.1; -.
CC FIR; A00654; TVWMD.
CC HSP; F11362; IFGK.
CC InterPro; IPR007110; Ig-like.
CC InterPro; IPR003598; Ig_c2.
CC InterPro; IPR003006; Ig_MHC.
CC InterPro; IPR000719; Prot_kinase.
CC InterPro; IPR001824; RtkinaseIII.
CC InterPro; IPR001245; Tyr_kinase.
CC Pfam; PF00047; Ig_3.
CC Pfam; PF00069; pkinase; 1.
CC ProDom; PD000001; Prot_kinase; 2.
CC SMART; SM00408; Igc2; 1.
CC SMART; SM00219; Tyrc; 1.
CC PROSITE; PS50835; IG LIKE; 3.
CC PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
CC PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
CC PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
CC PROSITE; PS00240; RECEPTOR_TYR_KIN_III; 1.
CC PolyProtein; Tyrosine-protein kinase; Oncogene; Transferase; Receptor;
KW Transmembrane; Glycoprotein; Phosphorylation; ATP-binding;
KW Immunoglobulin domain; Repeat.
FT DOMAIN 1 543 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 544 568 POTENTIAL.
FT DOMAIN 569 978 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 55 134 IG-LIKE C2-TYPE 1.
FT DOMAIN 141 231 IG-LIKE C2-TYPE 2.
FT DOMAIN 236 331 IG-LIKE C2-TYPE 3.
FT DOMAIN 333 431 IG-LIKE C2-TYPE 4.
FT DOMAIN 434 533 IG-LIKE C2-TYPE 5.
FT DOMAIN 613 942 PROTEIN KINASE.
FT NP_BIND 619 627 ATP (BY SIMILARITY).
FT BINDING 647 647 ATP (BY SIMILARITY).
FT DISULFID 76 118 POTENTIAL.
FT DISULFID 161 211 POTENTIAL.
FT DISULFID 258 312 POTENTIAL.
FT DISULFID 451 516 POTENTIAL.
FT MOD_RES 841 841 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT MOD_RES 973 973 PHOSPHORYLATION.
FT CARBOHYD 79 79 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 107 107 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 128 128 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 187 187 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 309 309 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 320 320 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 336 336 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 369 369 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 444 444 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 511 511 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 524 524 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CONFLICT 714 714 L -> P (IN REF. 1).
FT CONFLICT 971 978 ORTPPVAR -> RGPEL (IN REF. 1).
FT SEQUENCE 978 AA; 108491 MW; 4C7CAC4835185EBF CRC64;
SQ

```

Query Match	23.7%	Score 1249.5	DB 1	Length 978
Best Local Similarity	33.2%	Pred. No. 3.6e-76		
Matches	328	Conservative 162	Mismatches 348	Indels 149
		Gaps	32	
Qy	75	TVVTEAAVTEVAESGSIITLQVLAQPGDLSCU-----W---VFKHSSIGQCPHFDLQNRG	125	
Db	47	TAMHAQGPVPIQSPGPELVPEGTIVTLRCVNGSVWDGPISEHWNLDLDPSSII-----102		
Qy	126	IVSMAILNVTEQTAGEYLLHI-----QSERANYTVLFTVNRDRTOLYYL---RRPY-----173		
Db	103	---LTTNNATFQNTGY---HCTEPGNPRGGNATI-----HLVYKDPARPKVLAQE	148	
Qy	174	FRKMENODALL-CISEGVPBPTWVWLCSSHRECKEGBPAVVRKEKVLHELPGTDI--230		
Db	149	VTVLEGODALLPCL---LTPDALE-----AGVSLVRVGRPVLRTQNTYFSPWHGFTIHK	200	
Qy	231	-----RCCARNALGREGCTKL-----FTIDLNOAPQSTL-PQLFLKV-GEPLWIRCK	274	
Db	201	AKFTENHYQCSAR-VDRGTVTSMGIWKVKQDLSGPAITLLEPAELVRQGEAAQIVCS	259	
Qy	275	AIHVHFGFTWELEDKALEBGSFYFEMSTYSTNRTMIRILLAFVSSVGRNDTGYTTCSSS	334	
Db	260	ASNIDVNFDSVLRHGDTKL---TTSQSDPHDNRYQ-KVLTLNLDHVSFQDAGNYSCTAT	315	
Qy	335	K---HPSQSALVTILEKFINATSSQB-EYEIDPYEPCFSVRKAYPRIRC-TWI-----385		
Db	316	NAMGNHSGAMVFRVVEGAYSALNTSEQLLQVTVGKVDLQKVVEAYPGLESFNWTYILGP	375	
Qy	386	FSQASFCEQGRLEDGYSISKFCD---HKNKPGEYIFYAENDDAQFTKMTNIRKKPQ	441	
Db	376	FSDYQDKLDFVITKIDTVYTSLSPLRKSESGRYSLARNAGQNALTFELTLRYPE	435	
Qy	442	VLANASASQAS---CSSDGPPLPSWTWKKCSKSPNCTE---IPEGVWNKKNRNVFG	494	
Db	436	VRVTMTLINGSDDTLCEASGYPQSVTVVQCRSHTRCDSAGLVLEDSEVLVSQVFPY	495	
Qy	495	QWSSSTLNMSAAGKLLVKKCAVNSMGTSCTETFLNSPGFPFIQDNISFYATIGLC--552		
Db	496	EVIVHSLAIGTLESHNRTYECRAFNVSNGVSSOTWPISIGATPLPDELLFTPVLLTMS	555	
Qy	553	LPFIVVLVILCHKYKQRYESQLOMIQVTGPDNEFYVDFRDYEDYDLKWEFFRENL	611	
Db	556	IMALLLLELLLLLYKQKPYQVWKIIE---SYEGNSYTFIDPTQPLPYNEKWEFFRNLL	613	
Qy	612	EPGVIGSGAFGRVYNNATYAGISKTGVSIOAVKMLKEKADSCKEALMSLKMTHLGH	671	
Db	614	QFGKLTGTGAPKGVVETAPGLGKEDAVLKVAVKMLKSTAHADKEALMSLKMTHLGH	673	
Qy	672	HDNIVNLLGACTLSPVYLIFEYCCYGDLLNLYRSKRE-----KF	711	
Db	674	HENIVNLLGACTHGGPVLIVITEYCCYGDLLNLYRSKRE-----KF	733	
Qy	712	HRTWTEIFKEHNFSPYTFQAHNSNMPGSEVOLHPPDLQSGFNGNSIHSEDEIYEN	771	
Db	734	HLEKYYVRDRSGFS-----SQGVDYVEMRVSTSSNSDSFSEBDL-----774		
Qy	772	QKRLAEEEDLNLVITFEDLLCEAVOVAKGMFELEKFSCHVRDLAARNVLTGCKVVKIC	831	
Db	775	-----GKDRPLEURDLHFSSQVAGQWAFVASKNCHRDVAARNVLTSGRVAKIG	827	
Qy	832	DFGLARDILSDSSVYVRGNARLPVKWMAPESLFEGYITIKSDVMSYIGILLWEIFSLGNP	891	
Db	828	DFGLARDIMDSNVIVKGNARLPVKWMAPESLFDCVTVVQSDVMSYIGILLWEIFSLGNP	887	
Qy	892	YPGIPVDANFYKLIQSGFKMEQPPYATEGIYVVMQSCWAFDSRKRPPFNLTSLGQOLA	951	
Db	888	YPGILVNSKFYKLVKDGQYMAQAFAPAKNIYSIMQACWALPPTRRPTFQICSL--QKQ	945	
Qy	952	EABEACTRTSHLPKQAA-----PQOR	973	
Db	946	AQEDRRVFNNTLPPSSSSSSRLLRPWQR	972	

RESULT 7
 KPM5 HUMAN
 ID KPM5 HUMAN STANDARD; PRT; 972 AA.
 AC P07333;
 DT 01-APR-1988 (Rel. 07, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Macrophage colony stimulating factor I receptor precursor (CSF-1-R)
 DE (EC 2.7.1.112) (Fms proto-oncogene) (c-fms) (CD115 antigen).
 GN CSF1R OR FMS.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89239490; PubMed=2524025;
 RA Hampe A., Shamoon B.M., Gobet M., Sherr C.J., Galibert F.;
 RT "Nucleotide sequence and structural organization of the human FMS
 RT proto-oncogene.";
 RL Oncogene Res. 4:9-17(1989).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=86175013; PubMed=2421165;
 RA Coussens L., van Beveren C., Smith D., Chen E., Mitchell R.L.,
 RA Isacke C.M., Verma I.M., Ullrich A.;
 RT "Structural alteration of viral homologue of receptor proto-oncogene
 RT fms at carboxyl terminus.";
 RL Nature 320:277-280(1986).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Placenta;
 RX MEDLINE=9717923; PubMed=9027509;
 RA Andre C., Hampe A., Lachaux P., Martin E., Wang X.P., Manus V.,
 RA Hu W.X., Galibert F.;
 RT "Sequence analysis of two genomic regions containing the KIT and the
 RT FMS receptor tyrosine kinase genes.";
 RL Genomics 39:216-226(1997).
 RN [4]
 RP SEQUENCE OF 1-16 FROM N.A.
 RX MEDLINE=89261741; PubMed=2524648;
 RA Visvader J., Verma I.M.;
 RT "Differential transcription of exon 1 of the human c-fms gene in
 RT placental trophoblasts and monocytes.";
 RL Mol. Cell. Biol. 9:1336-1341(1989).
 RN [5]
 RP SEQUENCE OF 1-16 FROM N.A.
 RX MEDLINE=86218120; PubMed=3525854;
 RA Wheeler E.F., Roussel M.F., Hampe A., Walker M.H., Fried V.A.,
 RA Look A.T., Rettenmier C.W., Sherr C.J.;
 RT "The amino-terminal domain of the v-fms oncogene product includes a
 RT functional signal peptide that directs synthesis of a transforming
 RT glycoprotein in the absence of feline leukemia virus gag sequences.";
 RL J. Virol. 59:224-233(1986).
 CC -!- FUNCTION: THIS PROTEIN IS THE RECEPTOR FOR CSF-1, IT IS A PROTEIN
 CC -!- TYROSINE-KINASE TRANSMEMBRANE RECEPTOR.
 CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
 CC tyrosine phosphate.
 CC -!- SUBCELLULAR LOCATION: Type I membrane receptor.
 CC -!- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-
 CC PROTEIN KINASES.
 CC -!- SIMILARITY: Contains 5 immunoglobulin-like C2-type domains.
 CC -!- DATABASE: NAME=PROW; NOTE=PROW 1:13-21(2000);
 CC WWW="http://www.ncbi.nlm.nih.gov/prov/guide/1999017666.g.htm".
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; M25786; AAA58421.1; -;
 DR EMBL; U63963; AAB51696.1; -;
 DR EMBL; M14002; AAA35849.1; -;
 DR EMBL; X03663; CAA27300.1; -;
 DR PIR; S08123; TVHUMD.
 DR HSSP; P11362; IFGK.
 DR Genew; HGNC:2433; CSF1R.
 DR MIM; 164770; -;
 DR GO; GO:0005887; C:integral to plasma membrane; TAS.
 DR GO; GO:0005011; F:macrophage colony stimulating factor recept. .; TAS.
 DR GO; GO:0006960; P:antimicrobial humoral response (sensu Inver. .; TAS.
 DR GO; GO:0008283; P:cell proliferation; TAS.
 DR GO; GO:0007775; P:development; TAS.
 DR GO; GO:0007165; P:signal transduction; TAS.
 DR InterPro; IPR007110; Ig-like.
 DR InterPro; IPR003598; Ig_C2.
 DR InterPro; IPR003006; Ig_MHC.
 DR InterPro; IPR00719; Prot kinase.
 DR InterPro; IPR001824; RTKinaseIII.
 DR InterPro; IPR001245; Tyr_pkinase.
 DR Pfam; PF00047; Ig; 4.
 DR Pfam; PF00069; pkinase; 1.
 DR ProDom; PD000001; Prot kinase; 2.
 DR SMART; SM00408; IGC2; 1.
 DR SMART; SM00219; TyrKc; 1.
 DR PROSITE; PS00835; IG_LIKE; 3.
 DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
 DR PROSITE; PS00011; PROTEIN KINASE DOM; 1.
 DR PROSITE; PS00109; PROTEIN KINASE TVR; 1.
 DR PROSITE; PS00240; RECEPTOR_TYR_KIN_III; 1.
 KW Proto-oncogene; Tyrosine-protein kinase; Receptor; Transmembrane;
 KW Transferase; Glycoprotein; Phosphorylation; ATP-binding; Signal;
 KW Immunoglobulin domain; Repeat; Polymorphism.
 FT SIGNAL 1 19 POTENTIAL.
 FT CHAIN 20 972 MACROPHAGE COLONY STIMULATING FACTOR I
 FT RECEPTOR.
 FT DOMAIN 20 512 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 513 537 POTENTIAL.
 FT DOMAIN 538 972 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 21 104 IG-LIKE C2-TYPE 1.
 FT DOMAIN 107 197 IG-LIKE C2-TYPE 2.
 FT DOMAIN 203 230 IG-LIKE C2-TYPE 3.
 FT DOMAIN 299 339 IG-LIKE C2-TYPE 4.
 FT DOMAIN 402 502 IG-LIKE C2-TYPE 5.
 FT DOMAIN 582 910 PROTEIN KINASE.
 FT NP_BIND 588 596 ATP (BY SIMILARITY).
 FT BINDING 616 616 ATP (BY SIMILARITY).
 FT ACT_SITE 778 778 BY SIMILARITY.
 FT DISULFID 42 84 POTENTIAL.
 FT DISULFID 127 177 POTENTIAL.
 FT DISULFID 224 278 POTENTIAL.
 FT DISULFID 419 485 POTENTIAL.
 FT MOD_RES 708 708 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
 FT CARBOHYD 45 45 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 73 73 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 153 153 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 240 240 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 275 275 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 302 302 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 335 335 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 353 353 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 412 412 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 428 428 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 480 480 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT VARIANT 969 969 Y -> C (IN dbSNP:1801271).
 FT CONFLICT 54 54 /FTIC=VAR 011953.
 FT SEQUENCE 972 AA; 107983 MW; A8D99BE237573FE8 CRC64;
 SQ

Query Match 23.2%; Score 1221.5; DB 1; Length 972;
 Best Local Similarity 32.5%; Pred. No. 2.7e-74;
 Matches 330; Conservative 152; Mismatches 331; Indels 201; Gaps 33;

01-NOV-1995 (Rel. 32, Created)
 01-NOV-1995 (Rel. 32, Last sequence update)
 16-OCT-2001 (Rel. 40, Last annotation update)
 Mast/stem cell growth factor precursor (EC 2.7.1.112) (SCFR)
 DE (Proto-oncogene tyrosine-protein kinase Kit) (c-kit).
 GN Kit.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94215924; PubMed=7512939;
 RA Kubota T., Hikono H., Sasaki E., Sakurai M.;
 RT "Sequence of a bovine c-kit proto-oncogene cDNA";
 RL Gene 141:305-306(1994).
 CC -!- FUNCTION: THIS IS THE RECEPTOR FOR STEM CELL FACTOR (MAST CELL
 CC GROWTH FACTOR). IT HAS A TYROSINE-PROTEIN KINASE ACTIVITY.
 CC BINDING OF THE LIGANDS LEADS TO THE AUTOPHOSPHORYLATION KIT AND
 CC ITS ASSOCIATION WITH SUBSTRATES SUCH AS PHOSPHATIDYLINOSITOL
 CC 3-KINASE (PI3K).
 CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
 CC tyrosine phosphate.
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -!- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-
 CC PROTEIN KINASES.
 CC -!- SIMILARITY: Contains 5 immunoglobulin-like C2-type domains.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; D16680; BAA04084.1; -.
 DR HSP; P11362; 1FGK.
 DR InterPro; IPR007110; Ig-like.
 DR InterPro; IPR003599; Ig.
 DR InterPro; IPR003006; Ig_MHC.
 DR InterPro; IPR000719; Prot_kinase.
 DR InterPro; IPR001824; RTK_kinaseIII.
 DR InterPro; IPR001245; Tyr_kinase.
 DR Pfam; PF00047; ig; 1.
 DR Pfam; PF00069; pkinase; 1.
 DR ProDom; PD000001; Prot_kinase; 2.
 DR SMART; SM00409; IG; 3.
 DR SMART; SM00219; Tyrc; 1.
 DR PROSITE; PS00835; IG LIKE; 2.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
 DR PROSITE; PS00240; RECEPTOR_TYR_KIN_III; 1.
 KW Proto-oncogene; Tyrosine-protein kinase; Receptor; Transmembrane;
 KW Transferase; Glycoprotein; Phosphorylation; ATP-binding; Signal;
 KW Immunoglobulin domain.
 FT SIGNAL 1 22
 FT CHAIN 23 977
 FT DOMAIN 23 521
 FT TRANSMEM 522 544
 FT DOMAIN 545 977
 FT DOMAIN 590 938
 FT NP BIND 596 604
 FT BINDING 624 624
 FT ACT_SITE 793 793
 FT MOD_RES 824 824
 FT CARBOHYD 94 94
 FT CARBOHYD 130 130
 FT CARBOHYD 145 145
 FT CARBOHYD 284 284
 FT CARBOHYD 294 294

FT	CARBOHYD	301	301	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	321	321	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	353	353	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	368	368	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	401	401	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	464	464	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	487	487	N-LINKED (GLCNAC. . .)	(POTENTIAL).
SQ	SEQUENCE	977	AA; 109685	MW; 48271905088387EF	CRC64;
Query Match 23.1%; Score 1215; DB 1; Length 977;					
Best Local Similarity 32.8%; Pred. No. 7.4e-74;					
Matches 304; Conservative 141; Mismatches 298; Indels 184; Gaps 28;					
QY	134	VTE----	TOAGBYLLHIQSERANYTVLVTVNVDQ--LVTLRRPYPRKMNQDALLCI--	186	
DB	83	ITEKAEATNTGNYTC---TNKGGLSSSIYVFRDPEKFLIDLPLYGKEENDTLVRCLPT	139		
QY	187	-----SEGVPPEP--TVEWVLCSHRESCKEKGPAVVRKEEKLHELFGTDIRC--	232		
DB	140	DEVTNYSITGCEGKPLPKDLTFV-----ADPKAGITIRNVKREYHRL-----	CLH	185	
QY	233	CARNALGRE-CTKLFITIDLNAPOS-----TLPLFLKVGEPWLIRCKAIHVNHGFLT	285		
DB	186	CSANQKGKSMLSKFTLKVRAIKAVPVVSVSKTSYLLREGEFAVTCLIKOVSSSDSM	245		
QY	286	WELEKALEEGSYFEMSTYSTNRTWIRILLAFVSSVGRNDTGYTCCSSKHPQSALVTI	345		
DB	246	WIKENSQOTKQTKKNSWHQGDPSYLRQERLTATISSARVNDSGVFMCYANNTFGSANVT	305		
QY	346	LE---KGFINA--TSQOEYEIDPYKFCFVFKAYPR--IRCTWIFSQASFPCEQORGLD	400		
DB	306	LEVDKGFNIFPMNTTVFVNDGENVDLVVEAYEYKPKVHRQWIMYNNRT-----	355		
QY	401	GYSIKFCCHK-----NKGGEYIFYAENDDAQFTKMTLMIRKK	439		
DB	356	--STKWDYDPKSESNIRYVNELHLRLKGTGGTYTFHVSNSDVNSVTENVYNTK	413		
QY	440	POVLANASA--SQASCSDDGYPLPSMTWKCDKSPNCTEEI--PEGVMNKANRKFVGQW	496		
DB	414	PEILTHDLRVNMLQCVAAAGFPETIDMYFCPTGTEQRCSPVPGVDVQVQNSVSPFGKL	473		
QY	497	VSSSLNMEAGKGLLVKCCAYNSMGTS-----CETFLNSPGFPFFI	539		
DB	474	VYVSTIDDTFKHNGTVECRAYNDVGKSSAFNFAFGKSGKEQIHAHTLFT-----	527		
QY	540	QDNISFYATIGLCPPIVVLVILVILCHYKQPRYESQLQMI--QVTGPDNDNEYFYVDPDY	598		
DB	528	---IGFVIAAGLMCFVML-----TYKYLQKPMYEVQWKVVEING---NNYVIIDTQL	577		
QY	599	EYDLKWEFPRENLEFGKVLGSGAFGRVMNATAYIGTSIQVAVYKMLKEKADSCKEA	658		
DB	578	PYDHKWEFPNRNLSFGKTLGAGAGFKVVEATAYGLIKSDAAMTVAVKMLKPSAHLTEREA	637		
QY	659	LMSELKQWTHLGHNDIVNLLGACTLGGPYLYFEYCCYGDLLNLYRSKREKFRHTWEI	718		
DB	638	LMSELKVLSYLGNHNMIVNLLGACTTGGTTLVITEYCCYGDLLNLYRSKREKFRHTWEI	690		
QY	719	FKEHNPSSYPTFOAHSSNSMPGSRREVOLHPDLDLSGFNGNSIHSDEI-----EYENQ	772		
DB	691	-----ICSKQEDHAEVALYK-----NLHKSSESCNDSTNEYMDM	725		
QY	773	K-----FLAEEEEEDLN-----VLTFFEDLLCFAYQVAKGMEFLE	806		
DB	726	KPGVSVVVPFKADKRESARIGSYIERDVTPTAIMEDELALDLEDLLSFSQVAKGMAFLA	785		
QY	807	FKSCVHRDLAARNLVTHGKVKVICDFGLARDILSSSVVVRGNARLPVKWMAPESLFEG	866		
DB	786	SKNCIHRDLAARNILLTHGRITKICDFGLARDIKDNTSYVVKGNARLPVKWMAPESLFNC	845		
QY	867	IYTIKSDVMSYGLLWEIPEISLGVNYPGIPVDANFYKLIQSFGFKMQPPFYATGIIYPMQ	926		
DB	846	VYTFESDVMSYGIPLWELFSGLSSPYGPMVPVDSFKYMKIKGPRMLSPHAPAEYMDIMK	905		

QY 927 SCHAFDSRKRSPFNLTSLFLGCOLAEE 953
 Db 906 TCWDADPLKRPTEKQIVQLIEKQISES 932

RESULT 10
 KEMS_RAT
 ID KEMS_RAT STANDARD; PRT; 978 AA.
 AC Q00495;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Macrophage colony stimulating factor I receptor precursor (CSF-1-R)
 DE (EC 2.7.1.112) (Pms proto-oncogene) (c-fms).
 GN CSF1R OR CSFMR OR FMS.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Wistar; TISSUE=Skeletal muscle;
 RX MEDLINE=93001225; PubMed=1399227;
 RA Borycki A.G., Guiller M., Leibovitch M.P., Leibovitch S.A.;
 RT "Molecular cloning of CSF-1 receptor from rat myoblasts. Sequence
 analysis and regulation during myogenesis.";
 RL Growth Factors 6:209-218(1992).
 CC -!- FUNCTION: THIS PROTEIN IS THE RECEPTOR FOR CSF-1, IT IS A PROTEIN
 CC TYROSINE-KINASE TRANSMEMBRANE RECEPTOR.
 CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
 CC tyrosine phosphate.
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -!- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-
 CC PROTEIN KINASES.
 CC -!- SIMILARITY: Contains 5 immunoglobulin-like C2-type domains.

This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation
 at the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch)

EMBL; X61479; CAA43706.1; -
 PIR; I60321; S16385.
 HSP; P11362; IFGK.
 DR InterPro; IPR007110; Ig-like.
 DR InterPro; IPR003598; Ig C2.
 DR InterPro; IPR003006; Ig MHC.
 DR InterPro; IPR000719; Prot kinase.
 DR InterPro; IPR001824; RTKinaseIII.
 DR InterPro; IPR001245; Tyr_kinase.
 DR Pfam; PF00047; Ig; 4.
 DR Pfam; PF00069; pkinase; 1.
 DR ProDom; PD000001; Prot kinase; 2.
 DR SMART; SM00408; IGc2; 1.
 DR SMART; SM00219; Tyrcg; 1.
 DR PROSITE; PS00835; IG LIKE; 4.
 DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
 DR PROSITE; PS00011; PROTEIN KINASE DOM; 1.
 DR PROSITE; PS00109; PROTEIN KINASE TYR; 1.
 DR PROSITE; PS00240; RECEPTOR_TYR_KIN_III; 1.
 DR Proto-oncogene; Tyrosine-protein kinase; Receptor; Transmembrane;
 KW Transferase; Glycoprotein; Phosphorylation; ATP-binding; Signal;
 FT SIGNAL 1 19
 FT CHAIN 20 978
 FT MACROPHAGE COLONY STIMULATING FACTOR I
 FT RECEPTOR.
 FT DOMAIN 20 511
 FT EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 512 536
 FT POTENTIAL.
 FT DOMAIN 537 978
 FT CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 24 104
 FT IG-LIKE C2-TYPE 1.

FT DOMAIN 107 197
 FT DOMAIN 204 298
 FT DOMAIN 299 397
 FT DOMAIN 398 503
 FT DOMAIN 580 914
 FT NP BIND 586 594
 FT BINDING 614 614
 FT ACT SITE 776 776
 FT DISULFID 42 84
 FT DISULFID 127 177
 FT DISULFID 224 278
 FT DISULFID 417 483
 FT MOD RES 706 706
 FT CARBOHYD 45 45
 FT CARBOHYD 73 73
 FT CARBOHYD 302 302
 FT CARBOHYD 335 335
 FT CARBOHYD 389 389
 FT CARBOHYD 410 410
 FT CARBOHYD 449 449
 FT CARBOHYD 478 478
 FT CARBOHYD 491 491
 SQ SEQUENCE 978 AA; 109264 MW; 0A68456EF56BCTE3 CRC64;

Query Match 23.0%; Score 1210; DB 1; Length 978;
 Best Local Similarity 33.5%; Pred. No. 1.6e-73;
 Matches 318; Conservative 147; Mismatches 329; Indels 154; Gaps 31;

QY 78 EAATVEAEGSITLOVQLATPGDISLWFKHSSLGCPHPDLQNGRIVSMALNVET 137
 Db 36 ETVLRCSVNGSVEWD-----GPISPYWLTDPESPGS-----TLTRNATPK 77
 QY 138 QAGEY-LLIHQSERANYTLFTVNVVDLTOLYV-----LRRPYFRKMNQDALL-CIS 187
 Db 78 NTGTYRCTELEDPMAGSTI-----HLVYKDPAHSMNLLAQEVTVVGEQEAFLPCL- 128
 QY 188 EGVPEPTVWVLCSSHRESCKEKGPAVVRK-----BEKVLHELFGTDIRCC 233
 Db 129 -ITDPALK-----DSVSLMRGGGQVLRTVYFFSARWGPIIRKAVD-----DSNYVC 177
 QY 234 ARNALGRCTKL-FTIDLN-----QAPQSTL-PQLFLKV-GEPLWIRCKAIHVHNGFLTW 286
 Db 178 KTWNGRESTGTGLWLVKYNRVHPPEPQIKLEPSPKLVIRGEAAQVCSATNAEVEGNVL 237
 QY 287 ELEDKALE--EGSYFEMSTYSTNRTMIRILLAFVSSVGRNDTGYTC-SSSKHPQSALV 343
 Db 238 KRGDTKLEIPLNSDFQDNYKKVRAL-----SLNAVDQDAGIYSCVASNDVGTATM 291
 QY 344 T--ILEKGFINATSSQE--EYIDPYEKFCFVRFKAYPRIR-CTWIFSOASPPCEORGL 399
 Db 292 NFQVVEAYLNLITSEQLQEVSGDSLLITVHADAYPSIQHYNWYTLGPFFE--DQRKLE 350
 QY 400 -----DGYSISKFDH--KNKPEGYIFYAENDDAQFTKMTFLNIRKKPQVLANASQ 450
 Db 351 FITQRIAYRYTFKFLNLRVRKASEAGQYFLMAQNKAGWNLLTFELTLRYPPEVSVTWMPVN 410
 QY 451 AS----CSSDGYPLPSWTWKCKSDKSPNCTEEIIEGFWN-----KKAARKVFGQWVSSSTL 502
 Db 411 GSDVLFCDVSGYQPSVTVWMECRGHTDRCDQAQALQVWMDTHPEVLSPKDPKVIQSQL 470
 QY 503 NMSEAGKGLLVKCCAVNSMGTSCEITFLNSPGPPFFIODNISFYATIGLC---LPFIVVL 559
 Db 471 PIGTLKHNMTYFCKTNSVGNSSQYFRAVSLGQSKQLPDESFLTTPVAVACMSVMSLLVLL 530
 QY 560 IVLICHYKQKQRYEBSOLOMIQ-VTGPLDNEYFYDFRDYEDYDLKWEFFRENLEFFGKVLG 618
 Db 531 LLLLLYKQKPKYQVRWKIIEYEG---NSYTFIDTQLPYNEKEFFPRNNLQFGKTLG 587
 QY 619 SGAFGRVMNATYGISKTGVSIQVAVKMLKEKADSCKEALMSLKMTHLGHNDHNVNL 678
 Db 588 AGAFGKVETAFGLGKEDAVLVAVKMLKSTAHADKEALMSLKMTHLGHNDHNVNL 647
 QY 679 LGACTLSGPVYLIFEYCCYCDLLNLYRSKRE-----KPHRTWTEI 718

Db 509 -FAFEQIHPHTLPTLLIFGVIAAGMCIIMILTYLQKPMYEVQWKEEING--- 564
QY 588 NEXFYVDFRDYDLKWEFFRENLEFGKVGSGFGRVWNAATAYCISTGTGVISQVAKML 647
Db 565 NNVYIIDPTQLPYDHWKWEFFPNRLSFGKTLGAGAFKGVVEATAYGLIKSDAAMTVAVKML 624
QY 648 KERADSCKEKALMSKELKMTWHLGHHDNIIVNLGACTLSGGPVYLIFFEYCCYGDLENYLRSK 707
Db 625 KPSAHLTEREALMSKELKSVLGNHMIIVNLGACTVGGPTLVITEYCCYGDLENYLRK 684
QY 708 REXPHRTWTEIFKBNFSSPYTFOAHNSMPPGSGREVQLHPPLDQSGFNGNSIHSEDEI 767
Db 685 RDSFICKSQE---DHG-----EVALYK-----NLLHSKSS 712
QY 768 -----EYENQK-----RLAEEEDLN-----VLTFFDLCLFA 795
Db 713 CSSTNBYMDMKPGVSVYVPTKADKRSARIGSYIERDVTPIAMMEDDELALDLDLSFS 772
QY 796 YQVAKGMEFLEFKSVHRDLAARNVLVTHGKVKICDFGLARDILSDSSVYVVRGNARLPV 855
Db 773 YQVAKGNAFILAKNCICHRDLAARNILVTHGRITKICDFGLARDIKNSVYVVRGNARLPV 832
QY 856 KWMAPESLFEIGYITKISDVMSYGILLWEIFSLGVNPPGIPVDANFYKLIQSFGKMEQPPF 915
Db 833 KWMAPESIFNCVTFESDVMSYGIFLWELFSLGSSPYFGMPVDSKPYMKIKEGFRMLSP 892
QY 916 YATEGYFVMSQWAFDRSRKPPNLTSLFGCOLAEAEACIRTSIHLPKQAAPQORGG 975
Db 893 HAPAEYMDIMKTCDWDAPLKRPTFKQIVQLIEKQISDSTN-----HIYSNLAN----- 940
QY 976 LRAQSPORQVKIHRER 991
Db 941 -CSPNRPPVVDHSVR 955

RESULT 12

KIT_HUMAN
ID KIT_HUMAN STANDARD; PRT; 976 AA.
AC F10721;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Mast/stem cell growth factor receptor precursor (EC 2.7.1.112) (SCFR)
DE (Proto-oncogene tyrosine-protein kinase Kit) (c-kit) (CD117 antigen).
GN KIT.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Fetal brain, and Term placenta;
RX MEDLINE=88111521; PubMed=2448137;
RA Yarden Y., Kuang W.-J., Yang-Feng T., Coussens L., Munemitsu S.,
RA Dull T.J., Chen E., Schlessinger J., Francke U., Ullrich A.;
RT "Human proto-oncogene c-kit: a new cell surface receptor tyrosine
RT kinase for an unidentified ligand.";
RL EMBO J. 6:3341-3351(1987).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=93064697; PubMed=1279499;
RA Giebel L.B., Strunk K.M., Holmes S.A., Spritz R.A.;
RT "Organization and nucleotide sequence of the human KIT (mast/stem
RT cell growth factor receptor) proto-oncogene.";
RN Oncogene 7:2207-2217(1992).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=97179223; PubMed=9027509;
RA Andre C., Hampe A., Lachaume P., Martin E., Wang X.P., Manus V.,
RA Hu W.X., Galibert F.;
RT "Sequence analysis of two genomic regions containing the KIT and the
RT FMS receptor tyrosine kinase genes.";

Genomics 39:216-226(1997).

[4]
RN VARIANT LYS-583.
RX MEDLINE=92291284; PubMed=1376329;
RA Fleischman R.A.;
RT "Human piebald trait resulting from a dominant negative mutant allele
RT of the c-kit membrane receptor gene.";
RL J. Clin. Invest. 89:1713-1717(1992).
RN [5]
RN VARIANT LEU-584.
RX MEDLINE=92133600; PubMed=1370874;
RA Spritz R.A., Giebel L.B., Holmes S.A.;
RT "Dominant negative and loss of function mutations of the c-kit
RT (mast/stem cell growth factor receptor) proto-oncogene in human
RT piebaldism.";
RL Am. J. Hum. Genet. 50:261-269(1992).
RN [6]
RN VARIANT ARG-664.
RX MEDLINE=92020918; PubMed=1717985;
RA Giebel L.B., Spritz R.A.;
RT "Mutation of the KIT (mast/stem cell growth factor receptor)
RT protooncogene in human piebaldism.";
RL Proc. Natl. Acad. Sci. U.S.A. 88:8696-8699(1991).
RN [7]
RN VARIANT VAL-816.
RX MEDLINE=94013473; PubMed=7691885;
RA Furuta T., Tsujimura T., Tono T., Ikeda H., Kitayama H.,
RA Koshimizu U., Sugahara H., Butterfield J.H., Ashman L.K.,
RA Kanayama Y., Matsuzawa Y., Kitamura Y., Kanakura Y.;
RT "Identification of mutations in the coding sequence of the proto-
RT oncogene c-kit in a human mast cell leukemia cell line causing
RT ligand-independent activation of c-kit product.";
RL J. Clin. Invest. 92:1736-1744(1993).
RN [8]
RN VARIANTS PIEBALDISM GLY-791 AND VAL-812.
RX MEDLINE=93322624; PubMed=7687267;
RA Spritz R.A., Holmes S.A., Itin P., Kuester W.;
RT "Novel mutations of the KIT (mast/stem cell growth factor receptor)
RT proto-oncogene in human piebaldism.";
RL J. Invest. Dermatol. 101:22-25(1993).
RN [9]
RN VARIANT PIEBALDISM 893-GLU--PRO-896 DEL.
RX MEDLINE=96287384; PubMed=8680409;
RA Riva P., Milani N., Gandolfi P., Larizza L.;
RT "A 12-bp deletion (7818del12) in the c-kit protooncogene in a large
RT Italian kindred with piebaldism.";
RL Hum. Mutat. 6:343-345(1995).
RN [10]
RN VARIANT GIST VAL-559 DEL.
RX MEDLINE=98361155; PubMed=9697690;
RA Nishida T., Hirota S., Taniguchi M., Hashimoto K., Isozaki K.,
RA Nakamura H., Kanakura Y., Tanaka T., Takabayashi A., Matsuda H.,
RA Kitamura Y.;
RT "Familial gastrointestinal stromal tumours with germline mutation of
RT the KIT gene.";
RL Nat. Genet. 19:323-324(1998).
CC -1- FUNCTION: THIS IS THE RECEPTOR FOR STEM CELL FACTOR (MAST CELL
CC GROWTH FACTOR). IT HAS A TYROSINE-PROTEIN KINASE ACTIVITY.
CC BINDING OF THE LIGANDS LEADS TO THE AUTOPHOSPHORYLATION KIT AND
CC ITS ASSOCIATION WITH SUBSTRATES SUCH AS PHOSPHATIDYLINOSITOL
CC 3-KINASE (PI3K).
CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC tyrosine phosphate.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- DISEASE: DEFECTS IN KIT ARE THE CAUSE OF A DEVELOPMENTAL
CC ABNORMALITY KNOWN AS PIEBALDISM. AN AUTOSOMAL DOMINANT GENETIC
CC DISORDER OF PIGMENTATION CHARACTERIZED BY CONGENITAL PATCHES OF
CC WHITE SKIN AND HAIR THAT LACK MELANOCYTES.
CC -1- DISEASE: DEFECTS IN KIT ARE THE CAUSE OF A GASTROINTESTINAL
CC STROMAL TUMOR (GIST).
CC -1- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-
CC PROTEIN KINASES.
CC -1- SIMILARITY: Contains 5 immunoglobulin-like C2-type domains.

Db 662 IGGTLVITECCYGGDLNLFRRKDSFICSKQBDHAEALYKNLLHSSKSSCDSTNEY 721
Qy 744 VOLHPPDLQ---SGFNGNSHSDSEIYENOKLAEEDLNVLTPEDLCCFAYQVAK 800
Db 722 MDMPKGVSVVPTKADKERSVRISYIERDVTPTAIMEDEALDOL---EDLSFSYQVAK 778
Qy 801 GMFLEPKSCVHRDLAARNLVTHGKVKICDGLARDILSDSSYVVRGNARLPVKWAP 860
Db 779 GMAFLASKNCIHRDLAARNILLTHGRITKICDGLARDIKDNDYVVKGNARLPVKWAP 838
Qy 861 ESFLEGGYITKSDVMSGILLWEIYFSLGVNPGIPVDANFYKLIQSGFKMEQPFYATEG 920
Db 839 ESIFNCVYITPESDVMGIFLWELFSLGSPYDGPVDSKPYKMKIEGFRMLSPHAPAE 898
Qy 921 IYFVMSQAFDGRKPSFNLTSFLGCLQALAE 953
Db 899 MYDIMKTCWADPLKRTFKQVQLIEKQISES 931

RESULT 13
KIT CHICK
ID KIT CHICK STANDARD; PRT; 960 AA.
AC Q08156;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Mast/stem cell growth factor receptor precursor (EC 2.7.1.112) (SCFR)
DE (proto-oncogene tyrosine-protein kinase kit) (c-kit).
GN KIT.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauia; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=White leghorn; TISSUE=Brain;
RX MEDLINE=93292995; PubMed=7685729;
RA Sasaki E., Okamura H., Chikamune T., Kanai Y., Watanabe M.,
RA Naito M., Sakurai M.;
RT "Cloning and expression of the chicken c-kit proto-oncogene";
RL Gene 128:257-261(1993).
CC -!- FUNCTION: THIS IS THE RECEPTOR FOR STEM CELL FACTOR (MAST CELL
CC -!- GROWTH FACTOR). IT HAS A TYROSINE-PROTEIN KINASE ACTIVITY.
CC BINDING OF THE LIGANDS LEADS TO THE AUTOPHOSPHORYLATION KIT AND
CC ITS ASSOCIATION WITH SUBSTRATES SUCH AS PHOSPHATIDYLINOSITOL
CC 3-KINASE (PI3K).
CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC tyrosine phosphate.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- TISSUE SPECIFICITY: HIGH IN THE BRAIN AND TESTES AND ALSO
CC PRESENT IN THE BURSA OF FABRICUS, HEART, KIDNEY, LUNG, SPLEEN
CC THYMUS AND OVARY.
CC -!- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-
CC PROTEIN KINASES.
CC -!- SIMILARITY: Contains 5 immunoglobulin-like C2-type domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announcement/)
CC or send an email to license@isb-sib.ch)
CC
CC EMBL; D13225; BAA02506.1; -
CC PIR; JN0677; JN0677.
CC HSSP; P11362; 1FGK.
CC InterPro; IPR007110; Ig-like.
CC InterPro; IPR003598; Ig c2.
CC InterPro; IPR007119; Prot kinase.
CC InterPro; IPR001824; RTKinaseIII.

InterPro: IPR001245; Tyr_pkinase.
Pfam; PF00047; Ig; 3.
DR PFam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot kinase; 2.
DR SMART; SM00408; IGC2; 1.
DR SMART; SM00219; TyKc; 1.
DR PROSITE; PS00835; IG LIKE; 1.
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE; PS00011; PROTEIN KINASE DOM; 1.
DR PROSITE; PS00109; PROTEIN KINASE TYR; 1.
DR PROSITE; PS00240; RECEPTOR TYR KIN III; 1.
KW Proto-oncogene; Tyrosine-protein kinase; Receptor; Transmembrane;
KW Transferase; Glycoprotein; Phosphorylation; ATP-binding; Signal;
KW Immunoglobulin domain.
FT SIGNAL 1 24
FT CHAIN 25 960
FT DOMAIN 25 505
FT TRANSMEM 506 530
FT DOMAIN 531 960
FT DOMAIN 575 913
FT NP_BIND 581 589
FT BINDING 609 609
FT ACT_SITE 777 777
FT MOD_RES 808 808
FT CARBOHYD 76 76
FT CARBOHYD 135 135
FT CARBOHYD 149 149
FT CARBOHYD 269 269
FT CARBOHYD 286 286
FT CARBOHYD 306 306
FT CARBOHYD 318 318
FT CARBOHYD 338 338
FT CARBOHYD 343 343
FT CARBOHYD 356 356
FT CARBOHYD 453 453
FT CARBOHYD 469 469
SQ SEQUENCE 960 AA; 107311 MW; 0E93850527AB68F6 CRC64;

Query Match 22.4%; Score 1181; DB 1; Length 960;
Best Local Similarity 31.8%; Pred. No 1.4e-71;
Matches 310; Conservative 159; Mismatches 356; Indels 150; Gaps 30;

Qy 45 GSAGKPSRYMVRGSPEDLQCTPRRQSEGTVEATVEAESGITLQVLATPGDLSC 104
Db 24 GSPVHEESSLVVVKGEELKCN-----EEGPTVWFQNSDPSAKTR 65
Qy 105 L-----WPKHSSLCQPHFDLQNRGIVSMALNVTTQAGEYLLHIQSERANYTVLFTV 159
Db 66 ISNEKEWTKNATIRDIGECKSKG-----STVN-----SFYV 99
Qy 160 NVYRDTQLVYLRPRYFRKMNQDALLCISBGVPEPTVEWVLCSSHRESCKEKGPAVVRKEE 219
Db 100 FVKDENVLFLVDSLIIYGHKEDSLIVCLTDDP-VLNFTRKCDGKPLKNNMTFIPNPQK 158
Qy 220 ----KVLHELFQGTDIRCCAR-NALGRECTKLFTIDLNQAP-QSTLPQ-----FLKVG 266
Db 159 GIIIKNVQSRFKGCYQCLAKHNGVEKISEHIF--LNVRPVHKAFLPVITLSKYEYLLKEG 215
Qy 267 EPLMTRCAIHVHNGFGLTWELDKALBEGSYFEMSTYSTNTMIRILLAFVSSVGRNDT 326
Db 216 BEFEVTCITITDVSIVKASWISYKSAIVTSKSNLGDYGERK----LTLNIRSVGVNDS 271
Qy 327 GYYTCSSSKHP--SQSALVTI--LEKGFINATSSQE-EYEIDPYKFCFVSVPKAYPRIR 381
Db 272 GEFTC-QAENPFGKTNATVTLKALAKGVRVLFATWNTTIDINAGQGNLTVEYEPYK-- 328
Qy 382 CTWIFSQASPFCEQGLEGDYGSISKFDCHKNK-----PGEYIF 419
Db 329 -----PKEEVMVMYMETLQNSSDHYVKFTVGNNSYTSSELHLTRLKTEGIIYTF 378
Qy 420 YAENDDAQFTQNTLRKPKQVLANASAQ--ASCSDGYPLPFSWTWKCKSDKSPNCTE 477
Db 379 FVNSDASSSVTFNVVYKTKPEILTLDMLGNDILQCVATGFPAPTIYVYFCPGTQRCLD 438

between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
or send an email to license@isb-sib.ch)

CC EMBL; S76596; AAB33207.1; -
CC HSP; P11362; IFGK.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003006; Ig MHC.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR001824; RtkinaseII.
DR InterPro; IPR001245; Tyr_kinase.
DR Pfam; PF00047; Ig; 1.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 2.
DR SMART; SM00409; IG; 3.
DR SMART; SM00219; Tyrc; 1.
DR PROSITE; PS00835; IG LIKE; 2.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE; PS00240; RECEPTOR_TYR_KIN_III; 1.
KW Proto-oncogene; Tyrosine-protein kinase; Receptor; Transmembrane;
KW Transferase; Glycoprotein; Phosphorylation; ATP-binding; Signal;
KW Immunoglobulin domain.
FT SIGNAL 1 22 POTENTIAL.
FT CHAIN 23 978 MAST/STEM CELL GROWTH FACTOR RECEPTOR.
FT DOMAIN 23 521 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 522 544 POTENTIAL.
FT DOMAIN 545 978 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 590 938 PROTEIN KINASE.
FT NP BIND 596 604 ATP (BY SIMILARITY).
FT BINDING 624 624 ATP (BY SIMILARITY).
FT ACT_SITE 793 793 BY SIMILARITY.
FT MOD_RES 824 824 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT CARBOHYD 94 94 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 145 145 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 284 284 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 294 294 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 301 301 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 321 321 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 353 353 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 368 368 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 401 401 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 464 464 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 487 487 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 978 AA; 109449 MW; 6D45472E07440E6B CRC64;
Query Match 22.4%; Score 1178.5; DB 1; Length 978;
Best Local Similarity 31.6%; Pred. No. 2.1e-71;
Matches 313; Conservative 162; Mismatches 373; Indels 143; Gaps 30;
QY 36 CVLISHENSSAGKSSVVRMGSPEDLQCTPRROSEGVYEAATVEVAESSITLQVQ 95
DB 12 CVLLLLRLVQTGSSQPSA-----SPGWSLPSIH-PATSELIVSAGDEIRLL 57
QY 96 LATPGDLCLWPKHSLGCPHFQDNQNGIVSMAILNVTETQAGEYLLHIQSERANYTV 155
DB 58 CTDPGFVK--WTFE--TLG-QSSEITHNEWITEKA-----EATNTGNVTC---TNGGGLSS 105
QY 156 LFTVVRD--TOLVLRPRPKMENQDALLCI-----SEGVPEP----- 193
DB 106 SIYFVRDPAKFLVDLPYKGDHDTLVRCPITDPEVNTYSLRGCEKGLPKDLTFVTD 165
QY 194 -----TVEWVLCSSHR--ESCKEEGPAVVRKEKVLHELFGTDIRCCARNALGRECTKLF 246
DB 166 PKAGITIRNVKREYHRLCLHCSAD-----RKGSVLKSKFTLVRAAIR-----AVP 212
QY 247 TIDLNOAQSTLPQLFLKVGCEPLWIRCKALHVNHGFLGTWELEDKALEGSYFEMSYST 306

QY 478 E--IPEGWNKANKRV--FGQVSSSTLNMSAGKGLLVKCCAYNSMGTSCTETIFLS 532
DB 439 SPTISPMVKVSYTNSVPSFERILVESTVNAS-MFKSTGTICCEASSNGDKR-SSVFNF 496
QY 533 PGPFFIQDNISFYA-----TIGLCPPFIVLVILVILCHIKYKQFRYESQLQMI-QVTG 584
DB 497 A-----IKEQIRTHLTFTPLLIATFAGVAGLMCIIVMILVYIYLOKPKYEVQWVVEEING 551
QY 595 PLDNFYFVDYEDYDLKWEPPRENLEFGKVLGSGAGRVNATAYGSKTGVSQVAV 644
DB 552 ---NNYVIDPQLPYDHWEPFRLSPGKTLGAGAFKVEATAYGLFKSDAAMTAV 608
QY 645 KMLKEKADSCKEALMSLKMTHLGHHDNIYNLLGACTLSGFPVLIIFYCCYCGDLLNYL 704
DB 609 KMLKSAHLTEREALMSLKVLSYLGHNINYNLLGACTIGGTLVITEYCCYCGDLLNLF 668
QY 705 RSKREKF-----HRTWTE--IFKEHNPSSYPTQOA-HSNSMPSGSRVQLHPDLQLSGFN 757
DB 669 RKRDSFCPKHEEAHAAYENLLHQAEPTADAVNEMDMKPGVSYAVPPKADKRPVK 728
QY 758 GNSIHSEDEIEYENOKLAEEEDLNLVTFEDLLCFAYQVAKGMFLEFKSCVHRDLAA 817
DB 729 SGS-----YTDQVTLUSMEDDLALDVEDLSSFSYQVAKGMSFLASKNCIHRDLAA 780
QY 818 RNVLVTHGKWKICDFGLARDILSDSSVYVRGNARLPVKWMAPESEIFEGITYIKSDVWSY 877
DB 781 RNLLTHGRITKICDFGLARDIRNDSYVVKGNARLPVKWMAPESEIFCYTFESDVWSY 840
QY 878 GILLMEIFSLGNVPYGPVDFANFYKLIQSGFQMEQPFYATGFIYFMQSCWAFDSRKP 937
DB 841 GILLWELFSLGSPYGPMPVDFKXMKIEGYRMFSPGSPGPMYDMKSCWDADPLQRP 900
QY 938 SPNLTSLFGCLAE 952
DB 901 TPKQIVQLIEQQLSD 915
RESULT 14
KIT_FELCA STANDARD; PRT; 978 AA.
AC Q28889;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Mast/stem cell growth factor receptor precursor (EC 2.7.1.112) (SCFR)
DE (proto-oncogene tyrosine-protein kinase Kit) (c-kit).
GN KIT.
OS Felis silvestris catus (Cat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
OX NCBI_TaxID:9685;
RN SEQUENCE FROM N.A.
RP [1]
RC TISSUE=Head;
RX MEDLINE=95140426; PubMed=7530827;
RA Herbst R.; Munemitsu S.; Ullrich A.;
RT "Oncogenic activation of v-kit involves deletion of a putative
RT tyrosine-substrate interaction site."
RL Oncogene 10:369-379(1995).
CC -!- FUNCTION: THIS IS THE RECEPTOR FOR STEM CELL FACTOR (MAST CELL
CC GROWTH FACTOR). IT HAS A TYROSINE-PROTEIN KINASE ACTIVITY.
CC BINDING OF THE LIGANDS LEADS TO THE AUTOPHOSPHORYLATION KIT AND
CC ITS ASSOCIATION WITH SUBSTRATES SUCH AS PHOSPHATIDYLINOSITOL
CC 3-KINASE (PI3K).
CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC tyrosine phosphate.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-
CC PROTEIN KINASES.
CC -!- SIMILARITY: Contains 5 immunoglobulin-like C2-type domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2003, 07:20:28 ; Search time 83.4579 Seconds
(without alignments)
3067.272 Million cell updates/sec

Title: US-09-919-408a-2
Perfect score: 5264
Sequence: 1 MRALAQRSDRLLLLVLSV.....RGLRAQSPQVQKIHRRS 992

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL 23.*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phase.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*
15: sp_rvirus.*
16: sp_bacteriap.*
17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2099	39.9	406	11 Q8BR20	Q8br20 mus musculo
2	1268	24.1	976	13 Q8JPR5	Q8jf5 brachydanio
3	1268	24.1	976	13 Q9W755	Q9w755 brachydanio
4	1239	23.5	984	13 Q8AXC6	Q8axc6 fugu rubrip
5	1234.5	23.5	974	11 Q63702	Q63702 rattus ratt
6	1232.5	23.4	978	11 Q63116	Q63116 rattus norv
7	1225	23.3	979	11 Q8C8K9	Q8c8k9 mus musculo
8	1220	23.2	954	13 Q91909	Q91909 xenopus lae
9	1215.5	23.1	974	13 Q98SU3	Q98su3 danio dangi
10	1215.5	23.1	977	13 Q98SU1	Q98su1 danio nigro
11	1214	23.1	977	13 Q918N6	Q918n6 brachydanio
12	1209.5	23.0	979	6 Q9WN23	Q9wn23 canis fami
13	1208	22.9	964	6 Q97744	Q97744 sus scrofa
14	1207	22.9	964	6 Q9TQ01	Q9tcq1 sus scrofa
15	1207	22.9	964	6 Q9TQ00	Q9tcq0 sus scrofa
16	1205	22.9	978	6 Q9XS93	Q9xs93 canis fami

17	1204.5	22.9	977	13 Q98SU2	Q98su2 danio kerri
18	1201	22.8	948	6 Q9TDD7	Q9td7 trichosurus
19	1198.5	22.8	977	13 Q98SU4	Q98su4 danio albol
20	1197.5	22.7	992	13 Q9AXU0	Q9axu0 oncornychu
21	1194	22.7	972	4 Q99662	Q99662 homo sapien
22	1188.5	22.6	975	13 P79750	P79750 fugu rubrip
23	1173	22.3	945	6 Q77589	Q77589 equus cabal
24	1170.5	22.2	724	6 Q9MYN0	Q9myn0 bos taurus
25	1139.5	21.6	1062	13 Q8AXC7	Q8axc7 fugu rubrip
26	1136	21.6	1059	13 Q9DE49	Q9de49 brachydanio
27	1135.5	21.6	1078	13 Q8AXC8	Q8axc8 fugu rubrip
28	1126.5	21.4	1087	13 Q9PUF6	Q9puf6 gallus gall
29	1110	21.1	1097	11 Q8R406	Q8r406 rattus norv
30	1079	20.5	1106	4 Q8N5L4	Q8n5l4 homo sapien
31	1073	20.4	1019	13 Q8UVR8	Q8uvr8 fugu rubrip
32	1060	20.1	1048	13 P79749	P79749 fugu rubrip
33	1013.5	19.3	986	13 Q8UVR9	Q8uvr9 fugu rubrip
34	997.5	18.9	923	6 Q97745	Q97745 sus scrofa
35	980	18.6	1379	13 P79701	P79701 coturnix co
36	974	18.5	1327	13 Q8QHL3	Q8qhl3 gallus gall
37	961	18.3	1363	11 Q91ZT1	Q91zt1 rattus norv
38	957	18.2	1301	13 Q8UWU9	Q8uwu9 brachydanio
39	952.5	18.1	1302	13 Q8AXB3	Q8axb3 brachydanio
40	946	18.0	1173	13 Q9PTL0	Q9ptl0 brachydanio
41	939	17.8	563	11 Q925F7	Q925f7 rattus norv
42	937.5	17.8	1345	11 Q8VCD0	Q8vcd0 mus musculo
43	901.5	17.1	323	11 Q9EQ22	Q9eq22 rattus norv
44	877.5	16.7	323	11 Q9EQ24	Q9eq24 rattus norv
45	833	15.8	345	13 Q9PVU7	Q9pvu7 lethenteron

ALIGNMENTS

RESULT 1

Q8BR20 PRELIMINARY; PRT; 406 AA.

AC Q8BR20; (TREMREL. 23, Created)

DT 01-MAR-2003 (TREMREL. 23, Last sequence update)

DT 01-MAR-2003 (TREMREL. 23, Last sequence update)

DE FMS-like tyrosine kinase 3.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1] SEQUENCE FROM N.A.

RP STRAIN=C57BL/6J; TISSUE=Brain;

RX MEDLINE=22354683; PubMed=12466851;

RA The FANTOM Consortium,

RA the RIKEN Genome Exploration Research Group Phase I & II Team;

RT "Analysis of the mouse transcriptome based on functional annotation of

RT 60,770 full-length cDNAs."

RL Nature 420:563-573 (2002).

DR EMBL; AK045865; BAC32514.1; ..

SQ SEQUENCE 406 AA; 45834 MW; 155394D167D7D37 CRC64;

Query Match 39.9%; Score 2099; DB 11; Length 406;
Best Local Similarity 99.0%; Pred. No. 9e-176;
Matches 401; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 MRALAQRSDRLLLLVLSVILETVTNQDLPVTKCVLISHENNGSSAGKPSRYMRVGS 60
|||||

DB 1 MRALAQRSDRLLLLVLSVILETVTNQDLPVTKCVLISHENNGSSAGKPSRYMRVGS 60
|||||

QY 61 PEDLCQTPRROSEGTVEATVEAEGSITLQVLATPGDLISCLVFKHSLGQCPHFD 120
|||||

DB 61 PEDLCQTPRROSEGTVEATVEAEGSITLQVLATPGDLISCLVFKHSLGQCPHFD 120
|||||

QY 121 LQNRGIVSMALNVTETQAGYLLHIQSERANYTVLFTVNVVDTOLYVLRPFPRMENO 180
|||||

DB 121 LQNRGIVSMALNVTETQAGYLLHIQSERANYTVLFTVNVVDTOLYVLRPFPRMENO 180
|||||

```

QY 181 DALLCISGVPEPTVWVLCSSHRESCKEKGPAVVRKEKVLHFLGTDIRCCARNALGR 240
Db 181 DALLCISGVPEPTVWVLCSSHRESCKEKGPAVVRKEKVLHFLGTDIRCCARNALGR 240
QY 241 ECTKLFIDLNQAPQSTLPOLFLKVGEPWIRCKAIHVNHGFLTWELDKALEBSYFE 300
Db 241 ECTKLFIDLNQAPQSTLPOLFLKVGEPWIRCKAIHVNHGFLTWELDKALEBSYFE 300
QY 301 MSTYSTNRTMIRILLAFVSSVGRNDTGYTSCSSKHPQSQALVTILEKGFINATSSQBEY 360
Db 301 MSTYSTNRTMIRILLAFVSSVGRNDTGYTSCSSKHPQSQALVTILEKGFINATSSQBEY 360
QY 361 EIDPYKFCFVRKAYPRICRTWIFSOASPFCBOGLEDDGYSIS 405
Db 361 EIDPYKFCFVRKAYPRICRTWIFSOASPFCBOGLEDDGYSIS 405

RESULT 2
Q8JFR5
ID Q8JFR5 PRELIMINARY; PRT; 976 AA.
AC Q8JFR5
DT 01-OCT-2002 (TremBLrel. 22, Created)
DT 01-OCT-2002 (TremBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TremBLrel. 23, Last annotation update)
DE SI:d2116L04.1 (kit receptor (Tyrosine kinase)).
GN KIT.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN 1
RP SEQUENCE FROM N.A.
RA Pandian R.;
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL691516; CAD43458.1; -.
DR InterPro; IPR003599; IG.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003598; IG_c2.
DR InterPro; IPR003006; IG_MHC.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR001824; RtkinaseII.
DR InterPro; IPR002290; Ser_thr_kinase.
DR InterPro; IPR001245; Tyr_kinase.
DR Pfam; PF00047; Ig_3.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 2.
DR SMART; SM00409; IG; 4.
DR SMART; SM00408; IGc2; 3.
DR SMART; SM00220; S_TKc; 1.
DR SMART; SM00219; TyrKc; 1.
DR PROSITE; PS00835; IG_LIKE; 2.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE; PS00240; RECEPTOR_TYR_KIN_III; 1.
KW ATP-binding; Immunoglobulin domain; Kinase; Receptor; Transferase.
SQ SEQUENCE 976 AA; 10293 MW; A56921FA48DE58ED CRC64;

Query Match 24.1%; Score 1268; DB 13; Length 976;
Best Local Similarity 36.6%; Pred. No. 4.4e-102;
Matches 318; Conservative 138; Mismatches 302; Indels 110; Gaps 29;

QY 177 MENQDALLCISGVPEPTVWVLCSSHRESCKEKGPAV--VRKEKVLHFLGTDIRCCA 234
Db 138 MKNLNLQKC--DGQPLN-----SLRYSASLETGVSVEKVRKEFGCVCGTLDAAV 189
QY 235 RNALRECTKLTIDLNQAPQSTLP-----QLFLKVGEPWIRCKAIHVNHGFLTW 287
Db 190 KS--GR-----YQLTVRLVPDAPPITLQGPORVLLTQGEKLSLCSSTSNVNSDIATVWK 242
QY 288 LE---DKALEBSYFEMSTYSTNRTMIRILLAFVSSVGRNDTGYTSCSSKHPQSQA--- 341

```

```

Db 243 APNGVSVHONSHL-----LTEPIHTVHTAISLSVTVQDAGNYSCEAINEKGTAKPV 298
QY 342 LVTTLEKGFINAT--SQBEYEDPYKFCFVRKAYPRIRC-TWIFSOASPFCBOGLE 399
Db 299 WNIYKGFINITSVDNSTRVRAGESLSLRVVMNAYPKPHTFSWSYS----- 346
QY 400 DYSISKFCDH-----KPKGEYIIFYAENDDAQTFKMTFLNIRK 438
Db 347 -GVKLTNTDHTVTSRTHGNSYTSSELKVLRLKVSSEGIYTFSCLRDATIROTFEVHVIS 405
QY 439 KPOVLANASA--SQASCSGDVPLPSWTW-----KKCSKSPNCTEIPGCVNKKAN 489
Db 406 KPOIVSVEGPDIDGQVRCVAGGYPTPOIKWYICDLPHSRCSNLL-NATQE-EDVVTLTWT 463
QY 490 RKVFGQWVSSSTLNMSEAGKGLLVKCCAYNSMGTSCTIFLNSPGPPP---FIQDNISFY 546
Db 464 NPPFGKGAVERSLNITKNNYATLE--CVASANGEIVYTLFSISENTVPHELFTPLLGIV 521
QY 547 ATIGLCLPFIVLVILVILCHYKQFRIESQLQMTQVTOPLONEYFYVDFRVEYDLKWEF 606
Db 522 AAIVI-----LVLLILVLTQYKMQPKYQIQWKVIE--GIHGNNYVYIDPTQLPYDHQWEF 575
QY 607 PRENLEFGKVLGSGAFGRVMNATAYGISKTGVSIOQAVKMLKEKADSCKEALMSLKKM 666
Db 576 PRDKLRFKGLTGSAGFGKVEATAYGMSKADTVMIVAVKMLKPSAHATEKEALMSLKV 635
QY 667 TLHGHDNIIVNLLGACTLSGPVYLIFCYCCGDLNLNLSRKREKPHRTWTEIFKEHNFSS 726
Db 636 SYLGNHINIVNLLGACTVGGTLVITEYCCFGDLNLFRRRVFYFYT---TLGEDAYVR 692
QY 727 YPTFOAHNSNMPGSRVOLHPP-LDQLSGFNGNSIHSEDEIEYENQKRLAEEBEDLVN 785
Db 693 NVNQSEPNDSRNG--YMTKPSVLGILSENRRSLNKGD--SYSDSDAVSSEILOEDGLT 748
QY 786 LTFEDLLCFAYOVAKGMEFLFKSCVHRDLAARNVLVTHGKVVKICDFGLARDILSDSY 845
Db 749 LDTEDLISFSYQVAKGMDFLASKNCIHRDLAARNILLTQGRVAKICDFGLARDITTDSDY 808
QY 846 VVRGNARLPVKWMAPESLFEGITKSDVWSYGILLWEIFSLGVNVPYGPVVDANFYKLI 905
Db 809 VVKGARLPVKWMSPESEFECVYTFESDVWSYGILLWEIFSLGSSPYGMPVDSFKYKI 868
QY 906 QSGFKMEQPFYATEGIYFVMOSCAFDKSRKPSPNLTSFLGCLQAE-EEACITSTIHL 964
Db 869 KEGYRMESPEFSPEMYDIMHSCWDADVPKRPSEFKIVEKIEQQISDSTKHYLNFSSRL 928
QY 965 PKQAPQQRGGLRAQSPQORVKIHRERS 992
Db 929 P--AAP-----GPREESSH---VHRLNS 947

RESULT 3
Q9W755
ID Q9W755 PRELIMINARY; PRT; 976 AA.
AC Q9W755;
DT 01-NOV-1999 (TremBLrel. 12, Created)
DT 01-NOV-1999 (TremBLrel. 12, Last sequence update)
DT 01-MAR-2003 (TremBLrel. 23, Last annotation update)
DE Kit receptor tyrosine kinase.
GN KIT.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN 1
RP SEQUENCE FROM N.A.
RA Parichy D.M., Rawls J.F., Pratt S.J., Whitfield T.T., Johnson S.L.;
RT "Zebrafish sparse corresponds to an orthologue of c-kit and is
RT required for the morphogenesis of a subpopulation of melanocytes, but
RT is not essential for hematopoiesis or primordial germ cell

```

```

RT Development 126; 3425-3436(1999).
RL CC - SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).
CC CC - SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-
CC PROTEIN KINASES.
DR EMBL; AF153446; AAD41890.1; -.
DR HSP; P11362; 1FGK.
DR ZFIN; ZDB-GENE-980526-464; kit.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003598; Ig_c2.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR001824; RTKinaseIII.
DR InterPro; IPR001245; Tyr_kinase.
DR Pfam; PF00047; ig; 3.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 2.
DR SMART; SM00408; IGC2; 1.
DR SMART; SM00219; TyKc; 1.
DR PROSITE; PS50835; IG_LIKE; 2.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE; PS00240; RECEPTOR_TYR_KIN_III; 1.
KW ATP-binding; Glycoprotein; Immunoglobulin domain; Kinase;
KW Phosphorylation; Receptor; Transmembrane;
KW Tyrosine-protein kinase.
SQ SEQUENCE 976 AA; 109278 MW; A80AA01659C1A2DA CRC64;

Query Match 24.1%; Score 1268; DB 13; Length 976;
Best Local Similarity 36.6%; Pred. No. 4.4e-102;
Matches 318; Conservative 138; Mismatches 302; Indels 110; Gaps 29;

QY 177 MENQDALLCISGVPEPTVWVLCSSHRESKEEGPAV--VRKEEVLVHELFTDIRCCA 234
DQ 138 MKNLNLQKC--DQQLPN-----SLRYSASLETGVSQVKEPFEGCYVCGTLDAAIV 189
QY 235 RNALGRECTLFTIDLNQAPQSTLP-----QLFLKVGEPVWIRKAIHNVHFGLTWE 287
DQ 190 KS--GR-----YQLTVRLVPDAPPTLGGQPVRLVLTQGEKLSLSCSTSNVNSDIKVK 242
QY 288 LB---DKALEEGSFEMSTYSNTRMIRILLAPFVSVGRNDGYTCSSSKHPSQSA--- 341
DQ 243 APNGVNPVSHQSHL-----LTBPTHVTRAILSLSSVTWQDAGNSCEAINEKGTAKPV 298
QY 342 LVTILEKGFINATS--SQEYEIDPVEKFCFVRKAYPRIC--TWTFQASPPCEQGLE 399
DQ 299 WNIYIEKGFINTSDNSTRVRAGESLSLRVMNAYPKPHTFSWSYS----- 346
QY 400 DQYSISKFCDH-----KKNKPGYIFYAENDDAQFTKMTFLNIRK 438
DQ 347 -GVKLTNTTTHVITSRTHGNSYTSSELKVLRLKVSSEGIYTFSCNLRDATIRQTFVHVIS 405
QY 439 KPOVLNANASA--SQACSSSDGVPPLPSWTW-----KKCSDKSPNCTEIEIPGVWNNKAN 489
DQ 406 KPQIVSYEGPIDQVRCVAGPTPOIKWYCDLPHSRCNSLL--NATQB--EDWVTVTMT 463
QY 490 RRVFGQWSSSTLNKSEACKGLLVKCCAYNMGTSCTETIFLNSPGPPF---FIQDNISFY 546
DQ 464 NPPFGKGAVERSLNITKNYATLE--CVASANGEIVYTLFSENTPVHELTPLLIGTV 521
QY 547 ATIGCLPFIIVLILCHYKKQPRYESQLQMIQVTPGLDNEYFYVDPRDYEYDLKWEF 606
DQ 522 AAIVI---LVILILVLTVMQPKYQIQWKVIE--GIHGNVYVIDPTQLPYDHOWEF 575
QY 607 PRENLEFGVLGSGAGRVNNAATYAGISKTGVSIVQAVKMLKEKADSCKEALMSELKQM 666
DQ 576 PRDLKFRFGTKLGGAGFKVVEATYAGMSKADVTMTVAVKMLKPEPSAHATEKEALMSELKVL 635
QY 667 THLGHNDNVLNLLGACTLSGPPVILFEYCCYGDLLNLYRSKREKPHRTWTETIFKEHNFS 726
DQ 636 SYLGNHININLLGACTVGGPLVITEYCCFGDLLNFLRRRVYFYIT---TLGEDAYTR 692

```

```

727 YPTFOAHNSNMPGSGREVOLHPP--LDQLSGFNNGSIHSEDEIYENQKRLABEEEDLNV 785
DQ 693 NYVMQSEPNDSRNG--YMTMKSVLIGILSENRSLNKGD--SYSDSDAVSEILOEDCLT 748
QY 786 LTFEDLLCFAYQVAKMEFLFKSCVHRDLAARNVLVTHGVVKICDFGLARDILSDSSY 845
DQ 749 LUTEDLLSFYQVAKGMDFLASKNCIHRDLAARNILLTQGRVAKICDFGLARDITDTSNY 808
QY 846 VVRGNARLPVKWAPESLPEGIYTKSDVWSVIGILLWEIFSLGVNYPGIPVDANFYKLI 905
DQ 809 VYKGNARLPVKWSPESIFECVYTFESDVWSVIGILLWEIFSLGSSPYPGMPVDSKPYKI 868
QY 906 QSGFKMEQPFYATEGIYFVWQSCWAFDSRKRSFPNLTSLFLGCOLAEA--EACIRTSIHL 964
DQ 869 KEGYRMSEPFSEPMSEYIMHSCWDADPVKRSFSKIVKEIQOISDSTKHIIYLNFSRL 928
QY 965 PKQAAPOQRGGLRAQSPQOVKIHRRS 992
DQ 929 P--AAP-----GPRESSSH--VHRLNS 947

RESULT 4
Q8AXC6 PRELIMINARY; PRT; 984 AA.
AC Q8AXC6;
DT 01-MAR-2003 (Tremblrel. 23, Created)
DE 01-MAR-2003 (Tremblrel. 23, Last sequence update)
DE Kinase receptor c-kit.
GN C-KIT.
OS Fugu rubripes (Japanese pufferfish) (Takifugu rubripes).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Takifugu.
OX NCBI_TaxID=31033;
RN [1]
RP SEQUENCE FROM N.A.
RA Williams H., Brenner S., Venkatesh B.;
RT "Characterization of the Platelet-derived Growth Factor Receptor Alpha
RT and c-kit Genes in the Pufferfish Fugu rubripes.";
RL DNA Seq. 13:263-270(2002).
DR EMBL; AF456419; AAN87555.1; -.
KW Receptor.
SQ SEQUENCE 984 AA; 110732 MW; 27D1EDC837CB8295 CRC64;

Query Match 23.5%; Score 1239; DB 13; Length 984;
Best Local Similarity 32.4%; Pred. No. 1.6e-99;
Matches 334; Conservative 144; Mismatches 356; Indels 196; Gaps 34;

QY 60 SPEDLQCTPRQSGTVVEAATV--EVAESGSIITLQVQLATPGDLSCLVFKHSSLCQCP 117
DQ 38 APLELQCCQEKAKWQREERPKVGETKVDGKSTLYIPKAHPAHM----- 82
QY 118 HPDLQNGIVSMALNVTETQAGEY--LLHIQSERANYTLFTVNVVRDQLYVLRPYPR 175
DQ 83 -----GRYCLEETSOERAS-----IYIYVKDPD-----NPFK 111
QY 176 KM-----ENQDALL--CISEGVPEPTVWVLCSS-----HRESCKEKGPAVVRKEK 220
DQ 112 SMVSNILSREGDSASIPCLATPSLENLQKLTCKSSKALASGLHFSPLSQG--IIHNTQK 170
QY 221 VLH-----ELFGTDIRCCARNALGRECTLFTIDLNQAPQSTLPOLFLKVGEPVWIR 273
DQ 171 SYEGCVCTGRKLTENVRSHDHLTVRPVPVAPPVIMQAPK-----RVILRDESILYLC 226
QY 274 KAIHNVHGFGLTWEL---EDKALEEGSFEMSTYSNTRMIRILLAPFVSVGRNDGYT 330
DQ 227 NTTNVNGNIKWKWAPLGSQPAKVDGS---SRILTENFTQARSATLHIAAVRIQDTGRY 283
QY 331 CSSSKH---PSQALVTILEKGFINATS--SQEYEIDPVEKFCFVRKAYPRIR--CTWI 385
DQ 284 CBAENEKGVSTQSVLWDFEKGFMYSNPVNGTIQVRAGESLILLSVSI EAYPMRPSASWS 343

```



```

QY 386 FSQASPPCBQRLGLEDGYSISKFDH-----KNKPGYIFVAEND 424
Db 344 FM-----GRGLHNTSDHVITTRSHETYSSELKVLRLKMSGGVYTFQASNG 390
QY 425 DAQFTKMTFLNTRKKQVLANASA--SQASCSDDGYPLPSWTKKCDKSPNCTEII--- 479
Db 391 DASVNHFTTFIVISKDEIIVSHGEPDQVRCVAEGFPAPQITWYCEQPYARCQOVNAT 450
QY 480 --PEGVNNKANRKFQGVQVSSSTLNNSEAGKLLVKCCAYNSMGTSCETIFLNSPGPF-- 536
Db 451 QSEQNVITVTLSPFLGKTEVESRVNISR-GRSTLECYA-TVEGQAFTLFSISERTIS 508
QY 537 -----PFIONTSFYATIGLCLPFIIVLIVLICHKYKKQFYESQLQMTQVGPLNNEYF 591
Db 509 HOLFSPLLGSVS--AACILCL-----LIVLF-YKTMQPKYQIQWNVIE--GIHGNVY 559
QY 592 YVDFRDEYDLKWEFFRENLEKVLGSGAFGRVMNATAGISKTGVSIOVAVKMLKEKA 651
Db 560 YIDPTQLPYDQHWEPFRKNLRFQKTLGSGAFGKVEATAYGLANEDSMMTVAVKMLKSA 619
QY 652 DSCKEALMSLKMMTHLGHNDIVNLLGACTLGGPVYLIFECYCCYGDLLNLYRSKREKF 711
Db 620 HSTEKALMSLKVLIIYLNHINIVNLLGACTVGGPTLVITECCFGDLLNFLRRKRESF 679
QY 712 HRTWTEIFKEHNFSSYPTFOAHNSNMPGSRVQLHPLP--DOLSGF-----NGN 759
Db 680 -----ICFKLEEDCHY-----RNIMLOREMAGDSLNGYMTWRPSAAGKPS 720
QY 760 SHSEDEIEYQKRLAEBE-----EEDLNVTPEDLICFAYOVAKGHEFFLEKSCVHR 813
Db 721 SSSSEKRRSLREGSPVEEDSESEMEDSLSDTEDLSFSYQVAKGHEFFLTSKNCIHR 780
QY 814 DLAAARNVLTHGVKVKICDFGLARDILSDSSYVVRGNARLPVKWMAPESLFEGIYTIKSD 873
Db 781 DLAAARNILTOGRVAKICDFGLARDINTDSNVYVKGARLPVKWMAPESLFECVYTFESD 840
QY 874 VMSYGILLWEIFSLGVNYPGIPVDANFYKLIOSGFKMEOPFYATEGIVFVMSQWAFDS 933
Db 841 VMSYGILLWEIFSLGNSPYGMPVDAFKYLIKEGYRMDAPEFAPSEMYQIMESCWDADP 900
QY 934 RKRPSPNLTSFLGCOLAE-ABEACRTSIHLPKQAAPQORGLRAQS----- 980
Db 901 LNRPPPKVKVRIEQQLSDTTKHIYLNFSRVPMRGRRESSTHSMASQFNAGNSP 960
QY 981 PORQVKIHE 990
Db 961 PSRLPLHHE 970

```

RESULT 5

Q63702

ID Q63702 PRELIMINARY; PRT; 974 AA.

AC Q63702; (TEMBLrel. 01, Created)

DT 01-NOV-1996 (TEMBLrel. 01, Last sequence update)

DT 01-NOV-2003 (TEMBLrel. 23, Last annotation update)

DE C-kit receptor tyrosine kinase isoform.

OS Rattus rattus (Black rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

OX NCBI_TaxID=10117;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=BN/fmai.

RA Tsujimura T., Tono T., Yamazaki M., Nomura S., Kitamura Y.;

RT "Two isoforms of rat c-kit receptor tyrosine kinase."

RL Nucleic Acids Res. 0:0-0(0).

CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).

CC -1- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-

CC PROTEIN KINASES.

DR EMBL; X62491; CAA44354.1; -.

DR HSSP; P11362; 1FGK.

```

DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003598; Ig_c2.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR001824; RTKinaseII.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00047; ig; 1.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 2.
DR SMART; SM00408; IGC2; 1.
DR SMART; SM00219; TyKc; 1.
DR PROSITE; PS00835; IG LIKE; 3.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE; PS00240; RECEPTOR_TYR_KIN_III; 1.
KW ATP-binding; Glycoprotein; Immunoglobulin domain; Kinase;
KW Phosphorylation; Receptor; Transferase; Transmembrane;
KW Tyrosine-protein kinase.
SQ SEQUENCE 974 AA; 108955 MW; B9BFAD9BA0DA190D CRC64;

```

```

Query Match      23.5%; Score 1234.5; DB 11; Length 974;
Best Local Similarity 32.3%; Pred. No. 3.9e-99;
Matches 323; Conservative 176; Mismatches 340; Indels 161; Gaps 37;

```

```

QY 36 CVLISHENNGSAGKPPSYRMVRGSPEDLQCTPRROSEGTIVYEAAVEAEGSITLQVQ 95
Db 12 CVLLVLLRQGTGTSQPSA-----SPGEPSPPSIQPAQSELIVEAGD-TIRLT 57
QY 96 LATPGDGLCLWFKHSSLCQPHFDLQNRGIVSMALNV-TETQAGEYLLHIOSEIRANYT 154
Db 58 CTDPAFVK--WTFE-----ILDVRIENKQSEWIR--EKAETHT 92
QY 155 VLFT-----VNRVD-TQLVLRPRPRKMNQDALICISEGVPEPTV-ENVLC 200
Db 93 GKTCVSGSLRSSIVFVRDPAVLGLVGLFGKEDNALVRC---PLTDQPVSNYSLI 149
QY 201 SSHRESC-----KEEQPAVVRKEEVLHELFGTDIRCCARNALGRECTKLTIDLNQ 252
Db 150 ECDGKSLPTDLKFVNPVKAGITIKNVKRAYHRLC---IRCAAQREGKWMRSDFTLKVRA 206
QY 253 A-----POSTLPQI--FLKVGELWIRCKAIHVNHGFLTW-----ELEDKALEE 295
Db 207 AIKAIPVWVSPETSHLLKEGDTFTVICTIKDSTVSDSMWIKLNPQPSKAQKRNWHQ 266
QY 296 GSYFEMSTYSTNRTMIRILLAFVSSVGRNDTGYTCTSSSKHPQSQALVT---ILEKGIN 352
Db 267 GDF-----NYERQETLT-----ISSARVNDSGVFMCIYANNTFGSANVTTLKVKVEKGIN 316
QY 353 --ATSSQEEYIDPYEKFCSVRFKAYPR-IRCTWIFSQASFCEQGLE---DGYSIS 405
Db 317 IFPVKNTTVFVTDG-ENVDLVVEFEAYPKHQWIYMNRT--PTNRGEDYVKSDNQSN 373
QY 406 KFCDH-----KNKGEYIFVAENDDAOFTKMTFLNIRKKQVOLA--NASASQASCSDD 456
Db 374 RYVNELRLRLKGTGGTGYTFLVNSNDVASVTFDYVYVNTKPEILTYDRLMNGRLQCVAA 433
QY 457 GYPLPSMTWKCKSDKSPNCTEIEPE-GVNNKANRVFGQWSSSTLNMSEAGKGLLVKC 515
Db 434 GFPEPTIDWYFCGABQRCRTVPVVDVQIQNASVSPFGKLVVQSSIDSSVFRRHNGTVEC 493
QY 516 CAYNSMGTSCTEIFLNSPGFPFF--IQDN-----ISFYATIGLCLPFIIVLIVLICH- 565
Db 494 KASNAVGS--SAFFN---FAPKEIQPHLTFTPLLIGFVVVTAGL---MGIIVMVLAY 543
QY 566 KYKKQFRYESQLQMI-QVTGPDLDNEYVDYDRLDKWEPRENLEPKQVIGSGAFGR 624
Db 544 KYLQKPMYEVQWQVVEEING--NNYVYIDPTQLPYDHWKWEFFPRNRLSFGKTLGAGAFGK 600
QY 625 VMNATAYGISKTGVSIQVAVKMLKEKADSCKEALMSLKMMTHLGHNDIVNLLGACTL 684
Db 601 VVEATAYGLIKSDAAMTVAVKMLKPSAHLTERALMSSELKVLSTGLNHNMINVLLGACTV 660

```



```
QY 131 ILNVTTET-----QAGEYLHIIQSERANYTLVFTVNRDITQLYLVLPYPRKME 178
Db 69 -LNNSETDQFVLIKADLRHIGRYICTNTQENTSV--SLFVKDPAPELDFIDVTE 125
QY 179 NQDAL-LCISEGVPEPTVWVLCSSHRESCKEGPAVVRKEEVLHFLGTDIRC-----232
Db 126 GADTVGMCFFP---TDPMDIAI-----EKC--DGSPLPE-----NFTFTDIEAGITIK 169
QY 233 -----CARNALGR-ECTKLEFTIDLNAQPOSTLPQLFL-----KVGEPLWIRC 273
Db 170 TVQLAFDSCVCGNKGSTVKKSTTSIHVKVPKK-VPTVFLSKGRLQVKTEPEFVTC 228
QY 274 KAIHNVHGFLTW-ELEDKALEBSGFEMSTYTNRTMIRILLAFVSSVGRNDTGYVTC 332
Db 229 AVLDFSTVKAQWLDVKEGVTQKAFNRSSNVFYNLTLS-----DGVPSRSRFTQC 282
QY 333 SKHPSQ-----SALVTILEGFINATSOE-EYEDPYEKFCFVRPKAYPR-IRCTWIFS 387
Db 283 AENAIQGVNATFTLDVIDGVYVNLTVLENTTISVAGDNLVLKVYIDAYPHDPDGVWTF 342
QY 388 QASFPCEQRLGEGYISIKFCDHK-----KPEYIFYAENDDAQTKMP 432
Db 343 NETL-----LNTSDHYVATK--DEGNRYVSELHLRLKGTGKVYFTYTTNSDDASVSF 396
QY 433 TLNIRKKPOVLANASASQAS--CSSDGYPLPSWTWKKCDKSPNCTEEIPEGVWKK--A 488
Db 397 NIQVTRPEILLAEITSEGTLOCVATGFPVPAIQWYFCGSEQRCTDYPLSPVNEKFTQ 456
QY 489 NRKVFQGVSSSTLNMSEAGKLLVKKCCAYNSMGTS-----CETIFLNSPGP 535
Db 457 ENSSLGRIVVESTIDVNDLKNGTQCVASNEVESAYSVPFAIKELRTHLFT-----511
QY 536 PFIQDNTSFYATIGLCLPFVIVLVLICHVKYKQFRYESQLOMI-QVTGPLDNEYFYVD 594
Db 512 -PLL---IGFIAAGL-----MCIAVAVLMYKYLQKPKYEQWVBEING---NNVYID 560
QY 595 PRDYEDLKWEPPREMLRFGKVLGSAFGRVNATAYGSKTGVSIQAVKMLKEKADSC 654
Db 561 PTQLPYDNKWEPPRDLRCFGLKAGAGKVEATAYGLLKEDSRITVAVKMLKPSAHT 620
QY 655 EKEALMSELKMMTHLGHDNINVLGACLTSLGPVYLIFRYCCYGDLLNLYLRKREKFRHT 714
Db 621 EREALMSELKVLVYLGHGHNINVLGACITGVGFTLVITYCYCGDLLNLYLRKRDSF---677
QY 715 WTEIFKEHNFSYPTFOAHNSNMPGSRREVQHPPLDQLSGFNHNSIHSEDEIYENQK- 773
Db 678 -----ICPKFEDNS-----EALYKNL-----LNRDMGCEGMSYIDMKP 713
QY 774 -----RLAEEREDLNVLTPEDDLCPAYQVAKGMEFLFKSC 810
Db 714 AVSYVVPYTKDRSGSGDQDVSVSIPEEDDL-ALDTEDLINFYSYQVAGNVLASKNC 772
QY 811 VHRDLAARNVLVTHGKVKIKCDFGLARDILSDSYVVRGNARLPVKWMAPELSPGIVTI 870
Db 773 IHRDLAARNILLTHGRITIKCDFGLARDIRDSNVVVKGNARLPVKWMAPELSPGIVTF 832
QY 871 KSDVMSYGILLWEIPSLGVNYPGIPVDANFYKLIQSGFKMEQPPYATGIFYVMQSCWA 930
Db 833 ESDVMSYGILLWEIPSLGSSPYPRIPVDSKFKYKMIKDGVRMSPCAPLEMYEINRSCWN 892
QY 931 FDSRKRPSFPNLTSLGCOLAEAE 954
Db 893 SDPLKRPTFKQIVQMVQEQOLSDSK 916
```

RESULT 9

```
Q98SU3 ID Q98SU3 PRELIMINARY; PRT; 974 AA.
AC Q98SU3;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
```

```
DE Receptor tyrosine kinase Fms.
GN FMS.
OS Danio dangila.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=127599;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21359118; PubMed=11466528;
RA Parichy D.M., Johnson S.L.;
RT "Zebrafish hybrids suggest genetic mechanisms for pigment pattern
RT diversification in Danio.";
RL Dev. Genes Evol. 211:319-328 (2001).
CC -|- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).
CC -|- SIMILARITY: BELONGS TO THE CSP-1/PDGF RECEPTOR FAMILY OF TYROSINE-
CC PROTEIN KINASES.
DR EMBL; AF324479; AAK15301.1; -.
DR HSSP; P11362; 1FGK.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003598; Ig_C2.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR000719; Prot kinase.
DR InterPro; IPR001824; RTKinaseIII.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00047; Ig; 3.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot kinase; 2.
DR SMART; SM00408; IGC2; 1.
DR SMART; SM00219; TyrcK; 1.
DR PROSITE; PS00835; IG-LIKE; 3.
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE; PS00111; PROTEIN KINASE DOM; 1.
DR PROSITE; PS00109; PROTEIN KINASE TVR; 1.
DR PROSITE; PS00240; RECEPTOR TYR_KIN_III; 1.
KW ATP-binding; Glycoprotein; Immunoglobulin domain; Kinase;
KW Phosphorylation; Receptor; Transmembrane;
KW Tyrosine-protein kinase.
FT VARIANT 720 720 A -> T.
SQ SEQUENCE 974 AA; 109902 MW; DE7FD64FD0D78B37 CRC64;
```

```
Query Match 23.1%; Score 1215.5; DB 13; Length 974;
Best Local Similarity 34.6%; Pred. No. 1.8e-97;
Matches 339; Conservative 147; Mismatches 342; Indels 153; Gaps 37;

QY 54 YMWVSGSPEDLQCTPR-RQSECTVVEAATVEAEGSITLQVLA TPGLDLSCLWVFKHS 112
Db 13 FGQVGWSE-----PRILSSGAL--ADTDVILDSGS---PLQLCEGDPVTFL-----57
QY 113 LGCPHFDLQNRGI-----VSMALNVTTETQAGEY-LLHIQSERANYTVLFTVN 160
Db 58 ----PRLAKHKRYISKEVGKSRTPFRVETATVDT----GTAKIYINGNDSNLSSVHVFP 109
QY 161 VRDTP-LVY-----LRRPYFRKHENQDALLCISEGVPEPTVWVLCSSHRESCKEGPAV 214
Db 110 VRDSKLVFSPSTSLR--YVRK-EGEDLLPLCLLTPDPAT-DPIFRMDNGSAAPYGMNAT 165
QY 215 VRKEEKVL----HELFGTDIRCCARNALGRECTKLTIDLNAQPOSTLPQLFLK-----264
Db 166 FDPKGVLRNVHGFNADYICSAKIGGAEEKSKIFSINVIOQLRFP-PYVYLKNEYVK 224
QY 265 -VGEPLWIRCKAIHNVHGFLTWELEDKALEBSGFEMSTYTNRTMIRILLAFVSSVGR 323
Db 225 LVGERLQISCTNNPNFNSYNVTWTHSSRRLPAE--ERSTWEGDLAIESILT-IPSVOL 281
QY 324 NDTGYTCSSKHPSQSALVT---ILEKGF-----NATSSQEEYEIDPYEKFCFSV 372
Db 282 SDTGNIITCTGQNEAGANSSTTQLLVVDEPYIRLSPLKSSKLTNRDLSIEVSGDDVDLGV 341
QY 373 RFKAY-PRIRCTW---IFSQASFPCEQR--GLEDGYSISKFDCHKN--KPGYIFYAEND 424
Db 342 WIEAYPLPLSHKWETPTTSHNASLP-ENRFYNHNDRYEALLFLKRLNFEIIGYTLNVKNR 400
```


Q8WN23;
 01-MAR-2002 (TRENBLrel. 20, Created)
 01-MAR-2002 (TRENBLrel. 20, Last sequence update)
 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
 C-KIT.
 OS Canis familiaris (Dog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
 OX NCBI_TaxID=9615;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Zenne D., Yuzbasiyan-Gurkan V.;
 RL Submitted (NOV-2001) to the EMBL/GenBank/DDBJ databases.
 CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-
 CC PROTEIN KINASES
 DR EMBL; AF448148; AAL40833.1; --
 DR InterPro; IPR003599; IG.
 DR InterPro; IPR007110; IG-like.
 DR InterPro; IPR003006; IG_MHC.
 DR InterPro; IPR000719; Prot_Kinase.
 DR InterPro; IPR001824; RTKinaseIII.
 DR InterPro; IPR001245; Tyr_kinase.
 DR Pfam; PF00047; IG; 2.
 DR Pfam; PF00069; pkinase; 1.
 DR ProDom; PD000001; Prot_kinase; 2.
 DR SMART; SM00409; IG; 3.
 DR SMART; SM00219; TyrKG; 1.
 DR PROSITE; PS00835; IG-LIKE; 2.
 DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
 DR PROSITE; PS00111; PROTEIN KINASE DOM; 1.
 DR PROSITE; PS00109; PROTEIN KINASE TYR; 1.
 DR PROSITE; PS00240; RECEPTOR TYR KIN III; 1.
 DR ATP-binding; Glycoprotein; Immunoglobulin domain; Kinase;
 KW Phosphorylation; Receptor; Transferrase; Transmembrane;
 KW Tyrosine-protein kinase.
 SQ SEQUENCE 979 AA; 109753 MW; 46C30D5DBE8E33D3 CRC64;

Query Match 23.0%; Score 1209.5; DB 6; Length 979;
 Best Local Similarity 31.3%; Pred. No. 6.1e-97;
 Matches 332; Conservative 165; Mismatches 369; Indels 195; Gaps 35;

Qy 14 LLVLSVMILETVTNQDLVPVKLVLSHENSSAGKPSRYMRGSPEDLOCTPRROSE 73
 Db 11 LCVLLLLLLLGVGTGSSQPSV-----SPGERSL----- 38
 Qy 74 GTVVEAAVEVAESGSIITLQVLATPGDLSCLWVFKHSLGCPHFDLQNRGIVSMALIN 133
 Db 39 PSIIHAKSELIVSVGD-ELRLSCTDPGFVK--WTFE--TLG-QLNENTHNEWITEKA--- 89
 Qy 134 VTEQAGEVLLHIQSERANYTVLFTVVRD-TQLVLRPEYPRKMNQDALLCISEGVPE 192
 Db 90 ---EAGHTGNYTCNRDGLSRISIVFVRDPAKLFVLDLPYKGENDTLVRC----- 138
 Qy 193 PTVEWVLCSSHRESCKEPPAVRKEEVLHFLGTDIR-----C--CARNALGRE 241
 Db 139 PLTDEPVNYSLRGC--EGKPLPKDLTFVADPKAGITIRNVREYHRLCLHCSAQKGT 196
 Qy 242 -CTKLFTLDLNOAPG-----TLPOLFLKVEPLWIRKAIHVNHGFLTWLEDEKALE 294
 Db 197 VLSKKFTLVKRAAIRAVPVVSVKTSLSLLKEGEAFSVMCFIKDVSVFVSMWIKENSQOT 256
 Qy 295 EGSYFEMSTYSTNRTMIRILLAFVSGVNDGYTCSSSKHPSQSALVTILE---KGPI 351
 Db 257 NAQTSNWHHGGDFNFEROEKLIISARVNDSGVFMVCYANNFTGSAVNTTILEVDVKGFI 316
 Qy 352 NA-TSSQEEYIDPYEKFCFSVRKAYPR-IRCTMIFSQASF--PCEQRGLDEGYISIKF 407
 Db 317 NIFPMWSTITFVNDGENDVLIIVEYAYPKPEHQMIYNNRTITDRKWDYFKSDNESIRY 376
 Qy 408 CDH-----KNKPEYIPYAEANDDAQFTKMTFLNIRKPKQVLANASQA--SCSSDGY 458
 Db 377 YSELHLTRLKNGEGTYTFQVSNSDVNSSTFNVYVNTKPEILTHESLTNGMLQCVAAG 436

Qy 459 PLPSWTWKCKSDKSPNCTSEI--PEGVWVKANRKFVGQWVSSSTLMSBAGKGLLVKCA 517
 Db 437 PEPAVDWYFCGAEQRCVPIGPMVQMNSSLSPGKLVQSSIDYSAPKNGTVEGRA 496
 Qy 518 YNSMTGTCETIFLSPGFPFIQDN-----ISFYATIGLCPLFIVVLVLI 563
 Db 497 YNNVGRS--SAFFN----FAFKGSKQEQHPHTLFTPLLLIGFVIAAGM----MCIIVMIL 546
 Qy 564 CHYKKQFRYESQLQMI-QVTGPLDNEYFYVDFRDEYDLKWEEPRENLEFGKVLGSGAF 622
 Db 547 TYKLOKPMYEVQWKVBEING---NNVYIIDPTQLPYDHKWEFPNRLSFGKTLGAGAF 603
 Qy 623 GRVMNATAYGISTGVSQVAVKMLKEKADSCKEALMSLKMTHLGHHDHIVNLLGAC 682
 Db 604 GKVEATAYGLIKSDAAMTVAVKMLKPSAHLTEREALMSELKVLVSLGNHNVNLLGAC 663
 Qy 683 TUSGPVYLIFYCCYGDLLNLYRSKREKPHRTWTETFEKHNFSSTPTFOAHNSNSMPGSR 742
 Db 664 TVGGPTLVITEYCCYGDLLNFLRRKRDSPICSKQE---DHG----- 701
 Qy 743 EVQLHPPLDQLSGFNGNSIHSDEI-----EVENOK-----RLAEE 779
 Db 702 EVALYK-----NLHSESSCSDSTNEMDMKPGVSVVPTKADRRSARIGSYI 751
 Qy 780 EEDLN-----VLTPEDLLCFAYQYAKGMEFLFKSCVHRDLAARNVLVTHGKVKI 830
 Db 752 EEDVTPAIMEDDELALDLEDLLSFSYQYAKGWAFLASKNCIHRDLAARNILLTHGRITKI 811
 Qy 831 CDFGLARDILSDSSVYVGNARLPVKWMAPELSEFEGITYTIKSDVMSYIGILLWEIFSLGVN 890
 Db 812 CDFGLARDIKDSDVYVGNARLPVKWMAPELSEFEGITYTIKSDVMSYIGILLWEIFSLGVN 871
 Qy 891 PYGIPVDANFVKLIQSGFKMEQPFYATYEGIVFVMSQWAFDSRKEPSPNLTSLFGCOL 950
 Db 872 PYGMPVDSKFKYMIKEGFRMLSPHAPAEYDIMKTCDADPLKRPPTKQIVOLIEKQI 931
 Qy 951 AEAEACIRTSTHLPKQAPQORGGGLRAQSPQORQVKIHR 991
 Db 932 SDSTN-----HIYSNLAN-----CSPNPERPVVDHSVR 959

RESULT 13
 O97744
 ID O97744 PRELIMINARY; PRT; 964 AA.
 AC O97744;
 DT 01-MAY-1999 (TRENBLrel. 10, Created)
 DT 01-MAY-1999 (TRENBLrel. 10, Last sequence update)
 DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
 DE Mast/stem cell growth factor receptor (Fragment).
 GN KITI*0101.
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 OX NCBI_TaxID=9823;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Hampshire breed; TISSUE=Leukocyte;
 RX MEDLINE=98391767; PubMed=9724328;
 RA Marklund S., Kijas J., Rodriguez-Martinez H., Roennstrand L., Funa K.,
 RA Moller M., Edfors-Lilja I., Andersson L.;
 RT "Molecular evolution of the dominant white phenotype in the domestic
 pig";
 RL Genome Res. 8:826-833 (1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Hampshire breed; TISSUE=Leukocyte;
 RA Marklund S.;
 RL Thesis (1997), Department of Animal Breeding and Genetics,
 Swedish University of Agricultural Sciences, Uppsala, SWEDEN.
 CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-
 CC PROTEIN KINASES.

DR EMBL; AJ223228; CAAL1196.1; --
 DR HSSP; P11362; IFGK.
 DR InterPro; IPR003599; Ig.
 DR InterPro; IPR007110; Ig-like.
 DR InterPro; IPR003006; Ig MHC.
 DR InterPro; IPR000719; Prot kinase.
 DR InterPro; IPR001824; RTKinaselII.
 DR InterPro; IPR001245; Tyr_pkinase.
 DR Pfam; PF00047; ig; 1.
 DR Pfam; PF00069; pkinase; 1.
 DR ProDom; PD000001; Prot_kinase; 2.
 DR SMART; SM00409; IG; 3.
 DR SMART; SM00219; TyrcKc; 1.
 DR PROSITE; PS00835; IG-LIKE; 3.
 DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
 DR PROSITE; PS00011; PROTEIN KINASE DOM; 1.
 DR PROSITE; PS00109; PROTEIN KINASE TYR; 1.
 DR PROSITE; PS00240; RECEPTOR_TYR_KIN_III; 1.
 KW ATP-binding; Glycoprotein; Immunoglobulin domain; Kinase;
 KW Phosphorylation; Receptor; Transferase; Transmembrane;
 KW Tyrosine-protein kinase.
 FT NON_TER 964 964
 SQ SEQUENCE 964 AA; 108287 MW; 90E72EDFAB1359B7 CRC64;

Query Match 22.9%; Score 1208; DB 6; Length 964;
 Best Local Similarity 32.6%; Pred. No. 8.1e-97;
 Matches 319; Conservative 168; Mismatches 362; Indels 130; Gaps 32;

QY 38 LISHENNGSSAGKPSRYVRGSPEDLOCTPRQSGTVYEAATVEVAESGSLTQVLA 97
 Db 16 LLRLRVQTGSS--QPSV-----SPEL-----SPPSHPAKSELIVSAGD--EIRLFT 59
 QY 98 TFGDLSCLVWFVXHSGLGQCPHDLQNGIVSMALLNVTTQAGEYLLHTQSERA----- 151
 Db 60 DFG--SVKWTFF--TLG-----QLSENTHAEWIV-----EKAEAMNTG 93
 QY 152 NYTVL-----FTVNRVDTQ--LVYLRPRYFRKMNQDQ-----LCISE 188
 Db 94 NYTCTNEGGLSSIVYFVRDPKFLVDPPLVYGK--EDNDALVRCPLTDPVTNYSLTGCE 152
 QY 189 GVPEP-TVEWVLCSSHRESCKEAGPAVRKEVILHFGTDIRC--CARNALGRE-CTK 244
 Db 153 GKPLPKDITFV-----ADPKAGITIRNVKREYHRL-----CLHSANQGKSVLSK 198
 QY 245 LFTIDLNAPOS-----TLPOLFLKVGELWIRCKAIHNVHFGTLWELEDKALEBSY 298
 Db 199 KFTLKVRRAIAAPVAVSKASYLLREGEFAVMCLIKDVSSVDSMW--IRENSQTKAQV 257
 QY 299 FEMTYSYTRMIRILLAFVSSVGRNDTGYTTCSSSKHPSOSALVTILE--KGFINA-T 354
 Db 258 KRNSWHQGDENFLRQERLTISARVNDGVSFVFCYANNTFGSANVTITLVVDKGFNIPP 317
 QY 355 SSQEEYIDPYEKFCFSVRPKAYPR-IRCTWIFSQ--ASFCEQRLGSDYSISKFDCHK 411
 Db 318 MNMTTVFVNDGEDVDLIVEYEAYPEKPEHROWIYMNRTATDKWEDYPKSESNIRYSEL 377
 QY 412 N-----KPEGYFYAENDDAQTKMFTLNIRKPKQVLAN--ASASQASCSGSDYPLPS 462
 Db 378 HLTLKTEGTYTFLVSNADVNSVTFNVVNTKPEILTHDRLMGLMQLQCVAAGFPEPT 437
 QY 463 WTKKCKSDKSPNCTEEI--PEGVWNKKNRKFVGQWSSSTLNMSEAGGLLVKCCAYNSM 521
 Db 438 IDWVFCPTQRCSVPVGVVDVQVQNSVSPFGKLVTHSSIDYSAPKNGTVECRANDV 497
 QY 522 GTSCEIFLNSPGPPFPFIQONISFYATIGLCLPFIV-----VLIVLCHKYKKQFRYES 575
 Db 498 GK--SAFFN-----FAPKEQIHATLFTPLLIGFVIAAGMCIIVMLITYKYLQKPYEV 551
 QY 576 QLQMI-QVTGPLDNEYFYVDYEDLKWFFPRENLEFGKVLGSGAFGRMNATAYGIS 634
 Db 552 QWKVBEING--NNVYIDPTQLPYDHWKFFEPNRLSFGKTLGAGAFGKVVETAYGLI 608
 QY 635 KTVGSVQVAVMLKEKADSCKEALMSLKMWTHLGHHDHNVNLLGACTLSGPPYLIFEY 694

Db 609 KSDAAMTVAVKMLKPSAHLTEREALMSKLVSYLGNHNVNLLGACTIGGTTLVITY 668
 QY 695 CCYGDLLNYLRKREKFKHRTWTEIFKEHNFSSYPTFOAHSSNMPGSGREVQLHPLDOLS 754
 Db 669 CCYGDLLNFLRRKDSFICSQEDHAEALYKLLHKSKESSCSSTNMYMDKPGVSYV 728
 QY 755 GFNGNSHSEDEIYENQKRLAEEDDLNVLTPEDLCLCFAYQVAKMEFLFKSCVHRD 814
 Db 729 PTKADKRSARIGSYIERVDVTPAIMEDDELALDLEDLLSFSYQVAKGMAFLASKNCIHRD 788
 QY 815 LAARNVLVTHGVKVKICDFGLARDILSDSSVYVVGNAARLPVKWMAPESLFGIYTIKSDV 874
 Db 789 LAARNVLLTHGRITKICDFGLARDIKDINSYVVKGNARLPVKWMAPESLFGIYTIKSDV 848
 QY 875 WSYGILLWEIFSLGVNYPGIPVDANFYKLTOSQFKMEQPPYATYEGYVFWQSCWAPDSR 934
 Db 849 WSYGIFLWELFSLGSSPYPCMPVDSKPYKMKEGFRMLSPHEHAPAEIMDINKTCWDADPL 908
 QY 935 KRPSFNLTSFLGCQLAEA 953
 Db 909 KRPTFRQIVQLIEKQISES 927

RESULT 14

Q9TQ01 PRELIMINARY; PRT; 964 AA.
 AC Q9TQ01;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Mast/stem cell growth factor receptor (Fragment).
 GN KIT1*0201.
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 OX NCBI_TaxID=9823;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Yorkshire / Landrace cross; TISSUE=Leukocyte;
 RA Marklund S.;
 RL Thesis (1997), Department of Animal Breeding and Genetics,
 RL Swedish University of Agricultural Sciences, Uppsala, SWEDEN.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Yorkshire / Landrace cross; TISSUE=Leukocyte;
 RX MEDLINE=98391767; PubMed=9724328;
 RA Marklund S., Kijas J., Rodriguez-Martinez H., Roennstrand L., Funa K.,
 RA Moller W., Edfors-Lilja I., Andersson L.;
 RT "Molecular evolution of the dominant white phenotype in the domestic
 RT pig."
 RL Genome Res. 8:826-833(1998).
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-
 CC PROTEIN KINASES.
 DR EMBL; AJ223229; CAAL1197.1; --
 DR HSSP; P11362; IFGK.
 DR InterPro; IPR003599; Ig.
 DR InterPro; IPR007110; Ig-like.
 DR InterPro; IPR003006; Ig_MHC.
 DR InterPro; IPR000719; Prot kinase.
 DR InterPro; IPR001824; RTKinaselII.
 DR InterPro; IPR001245; Tyr_pkinase.
 DR Pfam; PF00047; ig; 1.
 DR Pfam; PF00069; pkinase; 1.
 DR ProDom; PD000001; Prot_kinase; 2.
 DR SMART; SM00409; IG; 3.
 DR SMART; SM00219; TyrcKc; 1.
 DR PROSITE; PS00835; IG-LIKE; 3.
 DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
 DR PROSITE; PS00011; PROTEIN KINASE DOM; 1.
 DR PROSITE; PS00109; PROTEIN KINASE TYR; 1.
 DR PROSITE; PS00240; RECEPTOR_TYR_KIN_III; 1.

KW ATP-binding; Glycoprotein; Immunoglobulin domain; Kinase;
KW Phosphorylation; Receptor; Transferase; Transmembrane;
KW Tyrosine-protein kinase.
FT NON TER 964 964
SQ. SEQUENCE 964 AA; 996C3C46201358A8 CRC64;

Query Match 22.9%; Score 1207; DB 6; Length 964;
Best Local Similarity 32.6%; Pred. No. 1e-96;
Matches 319; Conservative 168; Mismatches 362; Indels 130; Gaps 32;

QY 38 LISHENNGSSAGKPSYRMVGRGSPEDLOCTPRRQSEGTVEAATVEVAESGSLTQVOLA 97
DB 16 LLLRVQTGSS--QPSV-----SPEEL-----SPPSIHPAKSELIVSAGD-EIRLFCT 59
QY 98 TPGDLSCLVWFVXKSHSLGQCPHFQDLQNRGIVSMALNVTTQAGEYLLHIQSERA----- 151
DB 60 DPG--SVKWTFF--TLG-----QLSENTHAEWIV---EKAEMNTG 93
QY 152 NYTVL-----FTVVRDTQ-LVLRPRPFRKMNQDAL-----LCISE 188
DB 94 NYTCNEGGLSSIVFVRDPKFLVDPPLYGK-EDNDALVRCPLTDPVTNYSLTGCE 152
QY 189 GVPEP-TVEWVLCSSHRESKEEGPAVVRKEKVLHFGTDIRC--CARNALGRE-CTK 244
DB 153 GKPLPKDITFV-----ADPKAGITIRNVKREYHRL-----CLHGSANOGGKSVLSK 198
QY 245 LFTIDLNAPOS-----TLPOLFLKVGPEPLWIRCKAIHVNHVHFGTLWELEDKALEEGSY 298
DB 199 KFTLKVRAAIRAVPVAVSKASYLLREGGEFAVMCLIKDVSSVSDSMV-IRENSQTAKV 257
QY 299 FEMSTYSNRTWIRILLAFVSVGNRDTGYTCSSKHPQSALVTILE---KGFINA-T 354
DB 258 KRNSWHQGDFFLQERLTISSARVNDSGVFCYANVTTFLEVDVDFGINFP 317
QY 355 SSQBEYEIDPVKEKFCFVRKAYR-IRCTWLFQ--ASFPCQRLGEGYGISKPCDHK 411
DB 318 MNNTTVFVNDGVDVLIYEYEAIPKPHQWIMRATDKWEDYPKSENEINRYVEL 377
QY 412 N-----KPEYIFVABNDAAQFTKMTNIRKKPOVLAN--ASASQASCSDDGYPLPS 462
DB 378 HLRLKGTGEGGTYTFLVSNADVSSVTFNVVYNTKPEILTHRLNMGMLQCVAAFGPEPT 437
QY 463 WTKKCSKSPNCTEEI-PEGVWVKANKKRVQGWSSSTLNWSEAGKLLVKCCAYNSM 521
DB 438 IDWYPCGTEQRCSPVGPVQVQINSSVSPGKLVHSSIDYSAPKNGTVECRAYNDV 497
QY 522 GTSCTEILNSGPPFPQDNISFATIGLCPLPV-----VLIVLICHYKKQFRYES 575
DB 498 GKS--SAFFN-----FAFKEQIHAHTLFTPLLLGFVIAAGMCMIIIMLTLYLQKPMYEV 551
QY 576 QLQMI-QVTGPIQDNEYFYVDRDYDLKWEPPRENLEFGKLVGSGAFGRVNMATAYGIS 634
DB 552 QWKVVEENG---NNYVIDPTQLPYDHKWEFPNRLSPGKTLGAGAFKGVVEATAYGLI 608
QY 635 KTGVSIOAVKMLKEKADCEKALMSLKMTHLGHNDINVLNGLACTSLGSPVLYIFPY 694
DB 609 KSDAAMTAVAKMLKPSAHLTEREALMSLKLVSILGNHNIIVNLGACTIGGPTLVITEY 668
QY 695 CCYGLNLVLRSKREKFRHTWTETIPEKHNFSYPTFOAHNSNMPGSRREVQLHPDQLS 754
DB 669 CCYGLNLFLRRKDRDSFKQEDHAEALYKXNLHLSKSSSDSTNEYNDMKPGVSYV 728
QY 755 GFNGNSIHSDEIEYENQKRLABEEEDLNVLTFEDLLCFAYQVAKGMEFLFKSCVHRD 814
DB 729 PTKADKRSARIGSVIERDVTAIMEDDELALDLELLSFSQVAKGMAFLASKNCIHRD 788
QY 815 LAARNVLVTHGVKVICDFGLARDILSDSSVYVVRGNARLPVKWMAPESEFEGITYTKSDV 874
DB 789 LAARNILLTHGRITTKICDFGLARDIKNDSNVYVVGKGNARLPVKWMAPESEIFNCVYTFESDV 848
QY 875 WSYGILLWEIFSLGNVPYIGIPUDANFYKLIQSGFKMEQFPFATYEGIYFMQSCWAFDSR 934
DB 849 WSYGIFLWELFSLGSPFPGMPVDVDSKFKYKMKEGFRMLSPHAPVEMVIMTKCWDADPL 908

QY 935 KRPSFPNLTSFLGCOLAEE 953
DB 909 KRPTFKQIVQLIEKOISES 927

RESULT 15
Q9TQ00 PRELIMINARY; PRT; 964 AA.
AC Q9TQ00;
DT 01-MAY-2000 (TREMELrel. 13, Created)
DT 01-MAY-2000 (TREMELrel. 13, Last sequence update)
DT 01-MAR-2003 (TREMELrel. 23, Last annotation update)
DE Masc/stem cell growth factor receptor (fragment).
GN Kiti0202.
OS Sus scrofa (pig).
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RC STRAIN=Yorkshire / Landrace cross; TISSUE=Leukocyte;
RP SEQUENCE FROM N.A.
RA Marklund S.;
RL Thesis (1997), Department of Animal Breeding and Genetics,
RL Swedish University of Agricultural Sciences, Uppsala, SWEDEN.
RN [2]
RC STRAIN=Yorkshire / Landrace cross; TISSUE=Leukocyte;
RX MEDLINE=98391767; PubMed=9724328;
RA Marklund S., Kijas J., Rodriguez-Martinez H., Roennstrand L., Funa K.,
RA Moller M., Edfors-Lilja L., Andersson L.;
RT "Molecular evolution of the dominant white phenotype in the domestic
RT pig.";
RL Genome Res. 8:826-833(1998).
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-
CC PROTEIN KINASES.
DR EMBL; AJ223230; CAAL1198.1; -.
DR HSSP; F11362; IFGK.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003006; Ig MHC.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR001824; RtkinaseII.
DR Pfam; PF00047; Ig; 1.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 2.
DR SMART; SM00409; IG; 3.
DR SMART; SM00219; TyrKc; 1.
DR PROSITE; PS0835; IG LIKE; 3.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE; PS00240; RECEPTOR_TYR_KIN_III; 1.
KW ATP-binding; Glycoprotein; Immunoglobulin domain; Kinase;
KW Phosphorylation; Receptor; Transferase; Transmembrane;
KW Tyrosine-protein kinase.
FT NON TER 964 964
SQ SEQUENCE 964 AA; 108287 MW; 1B21292A962E9191 CRC64;

Query Match 22.9%; Score 1207; DB 6; Length 964;
Best Local Similarity 32.6%; Pred. No. 1e-96;
Matches 319; Conservative 168; Mismatches 362; Indels 130; Gaps 32;

QY 38 LISHENNGSSAGKPSYRMVGRGSPEDLOCTPRRQSEGTVEAATVEVAESGSLTQVOLA 97
DB 16 LLLRVQTGSS--QPSV-----SPEEL-----SPPSIHPAKSELIVSAGD-EIRLFCT 59
QY 98 TPGDLSCLVWFVXKSHSLGQCPHFQDLQNRGIVSMALNVTTQAGEYLLHIQSERA----- 151
DB 60 DPG--SVKWTFF--TLG-----QLSENTHAEWIV---EKAEMNTG 93

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2003, 07:13:34 ; Search time 62.4685 Seconds

(without alignments)

2520.578 Million cell updates/sec

Title: US-09-919-408A-2

Perfect score: 5264

Sequence: 1 MRALQSRDRLLLVLSV.....RGGLRAQSPQVKIHRS 992

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_19Jun03.*

1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5264	100.0	992	14	Murine flk-2. Mus
2	5264	100.0	992	16	Flk2 receptor prot
3	5264	100.0	992	16	Murine flk-2. Mus
4	5264	100.0	992	17	Human foetal liver
5	5264	100.0	992	18	Murine flk-2 recep
6	5264	100.0	992	20	Murine flk-2 prote
7	5264	100.0	992	23	Mouse receptor pr
8	5264	100.0	992	23	Murine receptor pr
9	5256	99.8	992	14	Murine flk-2 recep

10	5256	99.8	992	14	AAR31375	Murine flk-2. Mus
11	5239	99.5	992	13	AAR28038	Murine flk-2. Mus
12	5102	96.9	1000	16	AAR81868	Flk2/flt3 tyrosine
13	4436.5	84.3	993	16	AAR75961	Human STK-1. Homo
14	4429.5	84.1	993	16	AAR67816	Flk2 receptor prot
15	4429.5	84.1	993	16	AAR67536	Human flk-2. Homo
16	4429.5	84.1	993	17	AAR97419	Murine foetal live
17	4429.5	84.1	993	18	AAW19873	Human flk-2 recep
18	4429.5	84.1	993	20	AAV08617	Human flk-2 protei
19	4429.5	84.1	993	23	ABG70916	Human receptor pro
20	4429.5	84.1	993	23	AAE25819	Human receptor pro
21	4426.5	84.1	993	14	AAR37503	Human flk-2. Homo
22	4421.5	84.0	993	14	AAR44995	Human flk-2 recep
23	4420.5	84.0	993	16	AAR81869	Human flk2/flt3 ty
24	4417.5	83.9	1167	14	AAR31376	Human flk-2. Homo
25	4409.5	83.8	983	19	AAW63588	Human receptor typ
26	4408	83.7	986	19	AAW63587	Human receptor typ
27	4408	83.7	986	19	AAW63589	Human receptor typ
28	4404	83.7	994	19	AAW63586	Human receptor typ
29	2806.5	53.3	665	19	AAW63585	Human receptor typ
30	2373.5	45.1	749	19	AAW78002	Protein pMON32390.
31	2326.5	42.3	481	15	AAR47579	Soluble Flk-2. Mu
32	1461	27.8	307	23	ABB81198	Human FLT3 protein
33	1223.5	23.2	972	23	AAU11935	Colony stimulating
34	1223	23.2	975	22	AAE07144	Murine Kit/stem ce
35	1223	23.2	975	22	AAE07148	Mutant murine Kit/
36	1221.5	23.2	972	23	AAU79039	Human macrophage c
37	1221.5	23.2	972	23	AAU11936	Colony stimulating
38	1221.5	23.2	972	23	AAU11941	Colony stimulating
39	1221.5	23.2	972	23	AAU73585	Colony stimulating
40	1221.5	23.2	1055	22	ABG15479	Novel human diagno
41	1220.5	23.2	972	23	AAU11934	Colony stimulating
42	1220.5	23.2	972	23	AAU11937	Colony stimulating
43	1219.5	23.2	972	23	AAU11940	Colony stimulating
44	1216.5	23.1	972	23	AAU11938	Colony stimulating
45	1215	23.1	977	21	AAV51322	Bovine c-Kit bk-1

ALIGNMENTS

RESULT 1

ID AAR37502 standard; Protein; 992 AA.

XX AAR37502;

XX AC

XX 25-MAR-2003 (updated)

XX DT 19-OCT-1993 (first entry)

XX DE Murine flk-2.

XX Murine; receptor; protein; tyrosine kinase; pTK; flk-2; primitive;
XX hematopoietic cell; mature; family; conserved; region;
XX catalytic domain; c-kit; fetal liver kinase; flk; fetal; spleen;
XX thymus; adult; brain; bone marrow; multipotential; CFU-Blast colony;
XX hierarchy; transduction; T-lymphoid; lineage.

XX OS Mus musculus.

Key	Location/Qualifiers
FT Peptide	1..27
FT Protein	/note= "Hydrophobic leader sequence"
FT Domain	28..992
FT Region	/note= "Mature murine flk-2"
FT Domain	28..544
FT Region	/note= "Extracellular receptor domain"
FT Domain	545..564
FT Region	/note= "Transmembrane region"
FT Domain	565..992
FT Region	/note= "Intracellular catalytic domain"
FT Domain	618..623
FT Region	/note= "Catalytic sub-domain"

FT Domain 811..819 /note= "Catalytic sub-domain"
 FT Domain 832..834 /note= "Catalytic sub-domain"
 FT Domain 857..862 /note= "Catalytic sub-domain"
 FT Domain 872..878 /note= "Catalytic sub-domain"
 FT Region 736..812 /note= "flk-2 signature sequence"

WO9310136-A1.
 27-MAY-1993.
 16-NOV-1992; 92WO-US09893.
 15-NOV-1991; 91US-0793065.
 (UYPR-) UNIV PRINCETON.
 Lemischka IR;
 WPI; 1993-182479/22.
 N-PSDB; AAQ40914.

XX Totipotent haematopoietic stem cell receptors, their ligands and
 PT DNA sequences - for treating anaemia(s) and bone marrow damage
 PT due to e.g. cancer chemotherapy or radiotherapy

XX Claim 37; Fig 1a; 127pp; English.

XX This sequence represents the murine receptor protein tyrosine kinase
 CC (PTK), flk-2. The nucleic acid encoding this receptor is expressed
 CC in primitive hematopoietic cells and not in mature hematopoietic
 CC cells. Members of this family of PTK's can be recognised by the
 CC conserved amino acid regions in the catalytic domain. This family
 CC of PTK's also contains c-kit. These new receptors are termed fetal
 CC liver kinases (flk's) after the tissue in which they were discovered.
 CC flk-2 is also expressed in fetal spleen, fetal thymus, adult brain
 CC and adult bone marrow. flk-2 is expressed in individual multipotential
 CC CFU-Blast colonies capable of generating numerous multilineage colonies
 CC upon replating. It is likely therefore, that flk-2 is expressed in
 CC the entire primitive portion of the hematopoietic hierarchy. This is
 CC consistent with flk-2 being important in transducing putative self-
 CC renewal signals from the environment. flk-2 is the first receptor
 CC PTK known to be expressed in the T-lymphoid lineage.
 CC (Updated on 25-MAR-2003 to correct PN field.)

XX Sequence 992 AA;

Query Match 100.0%; Score 5264; DB 14; Length 992;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 992; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MRALAQRSDRLLLVLLVLSVILETVTNQDLPVTKVLI SHENNGSAGKPSVYRVGRS 60
 Db 1 MRALAQRSDRLLLVLLVLSVILETVTNQDLPVTKVLI SHENNGSAGKPSVYRVGRS 60
 Qy 61 PEDLOCTPRQSEGTVEAATVEAEGSITLQVLATPGDLSCLVWFKHSSLCQPHFD 120
 Db 61 PEDLOCTPRQSEGTVEAATVEAEGSITLQVLATPGDLSCLVWFKHSSLCQPHFD 120
 Qy 121 LQNRGIVSMALNTVETQAGYLLHIQSERANYTVLTVNVDRDTQLVLRPFRKMEQ 180
 Db 121 LQNRGIVSMALNTVETQAGYLLHIQSERANYTVLTVNVDRDTQLVLRPFRKMEQ 180
 Qy 181 DALLCISEGVPEPTVVEVWVLCSSHRESCKEKGPAVVRKEEKVLFHFGTDIRCCARNALGR 240
 Db 181 DALLCISEGVPEPTVVEVWVLCSSHRESCKEKGPAVVRKEEKVLFHFGTDIRCCARNALGR 240
 Qy 241 ECTKLFITIDLNOAQPSITLPQLFKVGPPLWIRCKAIHVNHFGLTWLEDKALEEGSYFE 300
 Db 241 ECTKLFITIDLNOAQPSITLPQLFKVGPPLWIRCKAIHVNHFGLTWLEDKALEEGSYFE 300

Db 241 ECTKLFITIDLNOAQPSITLPQLFKVGPPLWIRCKAIHVNHFGLTWLEDKALEEGSYFE 300
 Qy 301 MSTYSTNRTMIRILLAPFVSSVGRNDGYTCTSSSKHPSQSALVTILEKGFINATSSQEEY 360
 Db 301 MSTYSTNRTMIRILLAPFVSSVGRNDGYTCTSSSKHPSQSALVTILEKGFINATSSQEEY 360
 Qy 361 EIDPYEKFCFVSVPFKAYPRIRCTWIFSQASFPCEQGLEGDGYSTSKFCDHKNKPGEIFY 420
 Db 361 EIDPYEKFCFVSVPFKAYPRIRCTWIFSQASFPCEQGLEGDGYSTSKFCDHKNKPGEIFY 420
 Qy 421 AENDDAQFTKMFNTLNIRKKPOVLANASASQASCSSDGYPLPSWTWKCKSDKSPNCTBEIP 480
 Db 421 AENDDAQFTKMFNTLNIRKKPOVLANASASQASCSSDGYPLPSWTWKCKSDKSPNCTBEIP 480
 Qy 481 EGVNKKANRKFQGVQWVSSSTLNMSEAGKLLVKCCAYNSMGTSCTIFLNSPGPPFFIQ 540
 Db 481 EGVNKKANRKFQGVQWVSSSTLNMSEAGKLLVKCCAYNSMGTSCTIFLNSPGPPFFIQ 540
 Qy 541 DNISFYATIGLCIPFIVLVILVICHYKKQPRFYESQLOMTQVTGPDLDNEYFYVDFRDEY 600
 Db 541 DNISFYATIGLCIPFIVLVILVICHYKKQPRFYESQLOMTQVTGPDLDNEYFYVDFRDEY 600
 Qy 601 DLKWEPPRENLEFGKVLGSGAFGRVMNATAYGISTGVSIOAVKMLKEKADSCKEKALM 660
 Db 601 DLKWEPPRENLEFGKVLGSGAFGRVMNATAYGISTGVSIOAVKMLKEKADSCKEKALM 660
 Qy 661 SELKMTLHGHNDINVLGACTLSGPPVLIIFYCYYGDLNLNLYLRSKREKPHRTWTBIFK 720
 Db 661 SELKMTLHGHNDINVLGACTLSGPPVLIIFYCYYGDLNLNLYLRSKREKPHRTWTBIFK 720
 Qy 721 EHNFSYPTFOAHNSNSMPSGSRVQLHPPDLQSLSGFNNGSIHSEDEIEYENOKELAESEE 780
 Db 721 EHNFSYPTFOAHNSNSMPSGSRVQLHPPDLQSLSGFNNGSIHSEDEIEYENOKELAESEE 780
 Qy 781 EDNLVLTFFEDLLCFAYQVAKGMEFLBPKSCVHRDLAARNVLVTHGKVVKICDFGLARDIL 840
 Db 781 EDNLVLTFFEDLLCFAYQVAKGMEFLBPKSCVHRDLAARNVLVTHGKVVKICDFGLARDIL 840
 Qy 841 SDSYVVRGNARLPVKWMAPESLFEGITTIKSDVMSYGILLWEIFSGVNPYPGIPVDAN 900
 Db 841 SDSYVVRGNARLPVKWMAPESLFEGITTIKSDVMSYGILLWEIFSGVNPYPGIPVDAN 900
 Qy 901 FYKLIQSGFQWKEQFFYATGIIYFMQSCWAFDSKRPSPFNLTSLFLGCOLAEAEACIRT 960
 Db 901 FYKLIQSGFQWKEQFFYATGIIYFMQSCWAFDSKRPSPFNLTSLFLGCOLAEAEACIRT 960
 Qy 961 SIHLPKQAAPOQRGGLRAQSPQOVKIHRRS 992
 Db 961 SIHLPKQAAPOQRGGLRAQSPQOVKIHRRS 992

RESULT 2
 AAR67815
 ID AAR67815 standard; Protein; 992 AA.
 XX AAR67815;
 XX AC
 XX DT 25-MAR-2003 (updated)
 XX DT 16-AUG-1995 (first entry)
 XX DE Flk2 receptor protein-tyrosine-kinase.
 XX KW Mouse Flk2; receptor protein-tyrosine-kinase; primitive
 KW hematopoietic cell; fetal liver kinase; diagnostic ligand
 XX isolation; bone marrow disease therapy.
 OS Mus musculus.
 XX Key Location/Qualifiers
 FH Peptide 1..27
 FT Peptide /note= "signal peptide"
 FT Protein 28..992
 FT /note= "mature protein"

FT	Domain	28..544	/note= "extracellular receptor domain"
FT	Domain	545..564	/note= "transmembrane region"
FT	Domain	565..992	/note= "intracellular catalytic domain" }
FT	Domain	618..623	/note= "catalytic sub-domain"
FT	Domain	811..819	/note= "catalytic sub-domain"
FT	Domain	832..834	/note= "catalytic sub-domain"
FT	Domain	857..862	/note= "catalytic sub-domain"
FT	Domain	872..878	/note= "catalytic sub-domain"
FT	Region	709..785	/note= "signature sequence typical of Flk2"
XX			
PN	WO9500554-A2.		
XX			
PD	05-JAN-1995.		
XX			
PF	17-JUN-1994;	94WO-US06944.	
XX			
PR	18-JUN-1993;	93US-0080244.	
PR	21-JUN-1993;	93US-0081508.	
PR	23-NOV-1993;	93US-0157490.	
XX			
PA	(UYPR-) UNIV PRINCETON.		
XX			
PI	Lemischka IR;		
XX			
DR	WPI; 1995-052014/07.		
DR	N-PSDB; AAQ81012.		
XX			
PT	Ligand for receptor protein tyrosine kinase - useful for the		
PT	stimulation of primitive haematopoietic stem cells causing		
PT	proliferation and/or differentiation		
XX			
PS	Disclosure; Fig 1a; 131pp; English.		
XX			
CC	The sequence corresponds to a mouse Flk2 (fetal liver kinase)		
CC	receptor protein-tyrosine-kinase, which is expressed in primitive		
CC	hematopoietic cells but not in mature hematopoietic cells. The		
CC	protein is useful in isolation of receptor ligands, which have		
CC	applications in diagnosis of bone marrow disorders and in		
CC	stimulating proliferation and/or differentiation of primitive		
CC	hematopoietic stem cells.		
CC	(Updated on 25-MAR-2003 to correct PN field.)		
XX			
SQ	Sequence 992 AA;		
	Query Match 100.0%; Score 5264; DB 16; Length 992;		
	Best Local Similarity 100.0%; Pred. No. 0;		
	Matches 992; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
QY	1	MRALAQSDRRLLLLVLSWMILETVTNQDLPVIKCVLISHENNGSSAGKPSRYMRVGRS	60
Db	1	MRALAQSDRRLLLLVLSWMILETVTNQDLPVIKCVLISHENNGSSAGKPSRYMRVGRS	60
QY	61	PEDLOCTPRQSEGTVYEAAVVEAGSITLQVQLATPGDLSCLWVFKHSSLCQPHD	120
Db	61	PEDLOCTPRQSEGTVYEAAVVEAGSITLQVQLATPGDLSCLWVFKHSSLCQPHD	120
QY	121	LQNRGIVSMALNVTTQAGEYLLHIQSERANYTVLFTVNVDRDTQLVYLRPFYRKMEQ	180
Db	121	LQNRGIVSMALNVTTQAGEYLLHIQSERANYTVLFTVNVDRDTQLVYLRPFYRKMEQ	180
QY	181	DALLCITSEGPTEPTVEVWLCSHRESCKEGPAVVRKEEVLHFLFGTDTRCCARNALGR	240
Db	181	DALLCITSEGPTEPTVEVWLCSHRESCKEGPAVVRKEEVLHFLFGTDTRCCARNALGR	240

FT Region /label= Extracellular_receptor_domain
 FT 545..564
 FT /label= Transmembrane_region
 FT 565..992
 FT /label= Intracellular_catalytic_domain
 XX
 PN US3367057-A.
 XX
 PD 22-NOV-1994.
 XX
 PF 30-APR-1993; 93US-0055269.
 XX
 PR 02-APR-1991; 91US-0679666.
 PR 28-JUN-1991; 91US-0728913.
 PR 15-NOV-1991; 91US-0793065.
 PR 24-DEC-1991; 91US-0813593.
 PR 26-JUN-1992; 92US-0906397.
 PR 12-NOV-1992; 92US-0975049.
 PR 19-NOV-1992; 92US-0977451.
 PR 30-APR-1993; 93US-0055269.
 XX
 PA (UYPR-) UNIV PRINCETON.
 XX
 PI Lemischka IR;
 XX
 DR WPI; 1995-005894/01.
 DR N-PSDB; AAQ79068.
 XX
 PT Murine flk-2 receptor protein tyrosine kinase - used to stimulate
 PT proliferation and/or stimulation of primitive mammalian
 PT haematopoietic stem cells in vitro or in vivo.
 XX
 PS Claim 1; Fig. 1A-1F; 69pp; English.
 XX
 CC cDNAs encoding receptor protein tyrosine-kinases, mouse fetal liver
 CC kinase-2 (flk-2), human flk-2 and mouse flk-1 are given in AAQ79068-70,
 CC respectively, and the deduced amino acid sequences in AAR67535-37,
 CC respectively.
 CC (Updated on 25-MAR-2003 to correct PF field.)
 XX
 SQ Sequence 992 AA;
 Query Match 100.0%; Score 5264; DB 16; Length 992;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 992; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MRALQSRDRRLLLVLLVLSWMILETVNODLPVVKVLSHNNSSAGKPSYRMVRGS 60
 DB 1 MRALQSRDRRLLLVLLVLSWMILETVNODLPVVKVLSHNNSSAGKPSYRMVRGS 60
 QY 61 PEDLOCTPRRQSEGTVEAATVEAVERSGSITLQVQLATPGDLSCLVFKHSSILGCQPHFD 120
 DB 61 PEDLOCTPRRQSEGTVEAATVEAVERSGSITLQVQLATPGDLSCLVFKHSSILGCQPHFD 120
 QY 121 LQNRGIVSMALNVTTQAGEYLLHTQSERANTVLTNNVTRDTQYVLRPRFRKQWQ 180
 DB 121 LQNRGIVSMALNVTTQAGEYLLHTQSERANTVLTNNVTRDTQYVLRPRFRKQWQ 180
 QY 181 DALLCISEGVPEPTVEMVLCSSHRECKEGRPAVRKEKVLHELPGTDIRCCARNALGR 240
 DB 181 DALLCISEGVPEPTVEMVLCSSHRECKEGRPAVRKEKVLHELPGTDIRCCARNALGR 240
 QY 241 ECTKLFITDLNQAPQSTLPOLFLKVGEPMLWIRCKAIHVNHGFGLTWELEDKALEGSYPE 300
 DB 241 ECTKLFITDLNQAPQSTLPOLFLKVGEPMLWIRCKAIHVNHGFGLTWELEDKALEGSYPE 300
 QY 301 MSTYSTNRTMIRILLAFVSSVGRNDTGYTCCSSSKHPSQSALVTILEKGPINATSSQEEY 360
 DB 301 MSTYSTNRTMIRILLAFVSSVGRNDTGYTCCSSSKHPSQSALVTILEKGPINATSSQEEY 360
 QY 361 EIDPYEKFCFVSFRKAYPRIRCTWIIFSOAFCPEORGLDGYISIKFCDHKNPGEYIFY 420
 DB 361 EIDPYEKFCFVSFRKAYPRIRCTWIIFSOAFCPEORGLDGYISIKFCDHKNPGEYIFY 420

QY 421 AENDDAQFTKFTLNIRKKPQVLANASASQSSDGYPLPSWTWKKCSKSPNCTEETIP 480
 DB 421 AENDDAQFTKFTLNIRKKPQVLANASASQSSDGYPLPSWTWKKCSKSPNCTEETIP 480
 QY 481 EGVNKKANRKFVGQWVSSTLNMSEAGKGLLVKCCAYNSMGTSCETIFLNSGPPFFIQ 540
 DB 481 EGVNKKANRKFVGQWVSSTLNMSEAGKGLLVKCCAYNSMGTSCETIFLNSGPPFFIQ 540
 QY 541 DNISFYATIGLCLPFIVVLIVLICHKYKQFYESQLOMIQVTGPLDNEYFYVDFRDY 600
 DB 541 DNISFYATIGLCLPFIVVLIVLICHKYKQFYESQLOMIQVTGPLDNEYFYVDFRDY 600
 QY 601 DLKWEFPRENLEFGKVLGSGAFGRVWNTAYIGISKTGVSIOVAVKMLKEKADSCEKALM 660
 DB 601 DLKWEFPRENLEFGKVLGSGAFGRVWNTAYIGISKTGVSIOVAVKMLKEKADSCEKALM 660
 QY 661 SELKMTLGHHDNIVNLLGACTLGGPVYLI FEYCCYGDLLNLYRSKRKFHRTWTETIFK 720
 DB 661 SELKMTLGHHDNIVNLLGACTLGGPVYLI FEYCCYGDLLNLYRSKRKFHRTWTETIFK 720
 QY 721 EHNFSYPTFOAHSNSMPSGSRVQLHPPLDOLSGFNNGSIHSEDEI EYENQKRLAE 780
 DB 721 EHNFSYPTFOAHSNSMPSGSRVQLHPPLDOLSGFNNGSIHSEDEI EYENQKRLAE 780
 QY 781 EDLNVLTTPEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLTGHGVKVICDFGLARDIL 840
 DB 781 EDLNVLTTPEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLTGHGVKVICDFGLARDIL 840
 QY 841 SDSSVYVRGNARLPVKWMAPELSLFGIYTIKSDVMSYGLLWEI FSLGVNYPYGPVPDAN 900
 DB 841 SDSSVYVRGNARLPVKWMAPELSLFGIYTIKSDVMSYGLLWEI FSLGVNYPYGPVPDAN 900
 QY 901 FYKLQSGFKMEQPFYATGIVFMQSCWAFDSRKRSPFNLTSLFLGCQLAEAEACIRT 960
 DB 901 FYKLQSGFKMEQPFYATGIVFMQSCWAFDSRKRSPFNLTSLFLGCQLAEAEACIRT 960
 QY 961 SIHLPKQAAPOQGGRLAQSPOQVKIHRERS 992
 DB 961 SIHLPKQAAPOQGGRLAQSPOQVKIHRERS 992
 RESULT 4
 AAR97418
 ID AAR97418 standard; Protein; 992 AA.
 XX
 AC AAR97418;
 XX
 DT 25-MAR-2003 (updated)
 DT 11-DEC-1996 (first entry)
 XX
 DE Human foetal liver kinase 2.
 XX
 KW Human; foetal liver kinase 2; flk-2; protein tyrosine kinase;
 KW monoclonal; antibody; extracellular domain; receptor assay;
 KW haematopoietic stem cell; ligand; stimulation; proliferation;
 KW differentiation; treatment; anaemia; bone marrow damage;
 KW cancer chemotherapy; radiation.
 XX
 OS Homo sapiens.
 XX
 FH Key
 FT Peptide 1..27 /label= sig_peptide
 FT Peptide 28..992 /label= mat_peptide
 FT Domain 28..543 /label= extracellular_domain
 FT Domain 544..563 /label= transmembrane_domain
 FT Domain 564..992 /label= intracellular_domain
 FT
 XX

PN US548065-A.
 XX 20-AUG-1996.
 XX 31-OCT-1994; 94US-0252517.
 XX 19-NOV-1992; 92US-0977451.
 PR 02-APR-1991; 91US-0679666.
 PR 28-JUN-1991; 91US-0728913.
 PR 15-NOV-1991; 91US-0793065.
 PR 24-DEC-1991; 91US-0813593.
 PR 26-JUN-1992; 92US-0906397.
 PR 12-NOV-1992; 92US-0975049.
 PR 30-APR-1993; 93US-0055269.
 PR 31-OCT-1994; 94US-0252517.
 XX (UYPR-) UNIV PRINCETON.
 PA Lemischka IR;
 XX WPI; 1996-392678/39.
 DR N-PSDB; AAT38733.
 DR
 XX Anti-foetal liver kinase 2 (flk-2) antibodies - useful in assays,
 PT for isolating haematopoietic stem cells expressing receptor and for
 PT obtaining ligands
 XX
 PS Claim 1; Columns 27-34; 50pp; English.
 XX
 CC The present sequence is human foetal liver kinase 2 (flk-2), a
 CC protein tyrosine kinase. Isolated antibodies, pref. monoclonal,
 CC raised against the extracellular portion of flk-2 can be used to
 CC assay for flk receptors on the surface of primitive haematopoietic
 CC stem cells, and to isolate positive cells. The antibodies can also
 CC be used as, or to obtain ligands, which stimulate the proliferation
 CC and/or differentiation of stem cells. The ligands can be used, e.g.
 CC for treating anaemia, or bone marrow damage resulting from cancer
 CC chemotherapy, or radiation.
 CC (Updated on 25-MAR-2003 to correct PF field.)
 XX
 SQ Sequence 992 AA;
 Query Match 100.0%; Score 5264; DB 17; Length 992;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 992; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MRALQSRDRLLLVLSVMILEVTNQDLPVVKCVLISHENNGSSAGKPSYRMVRS 60
 Db 1 MRALQSRDRLLLVLSVMILEVTNQDLPVVKCVLISHENNGSSAGKPSYRMVRS 60
 QY 61 PEDLOCTPRQSEGTVEAATVEABSGSITLQVQLATPGDLSCLVFKHSLGCGPHFD 120
 Db 61 PEDLOCTPRQSEGTVEAATVEABSGSITLQVQLATPGDLSCLVFKHSLGCGPHFD 120
 QY 121 LQNRGIVSMALNVTTQAGEYLLHIQSERANYTVLTVNVRTQLYVLRPPYRKMENQ 180
 Db 121 LQNRGIVSMALNVTTQAGEYLLHIQSERANYTVLTVNVRTQLYVLRPPYRKMENQ 180
 QY 181 DALLCISEGVPEPTVEVWLCSSHRESCKEEGPAVVRKEKVLHELFGTDIRCARNALGR 240
 Db 181 DALLCISEGVPEPTVEVWLCSSHRESCKEEGPAVVRKEKVLHELFGTDIRCARNALGR 240
 QY 241 ECTKLTIDLNQAPOSTLPQLFKVGEPLWIRCKAIHNVNHGFLTWLEDKALEGSEYPE 300
 Db 241 ECTKLTIDLNQAPOSTLPQLFKVGEPLWIRCKAIHNVNHGFLTWLEDKALEGSEYPE 300
 QY 301 MSTYGTNRTMIRILLAFVSSVGRNDTGYTCSSKHPSQSALVTILEKGFINATSSQEEY 360
 Db 301 MSTYGTNRTMIRILLAFVSSVGRNDTGYTCSSKHPSQSALVTILEKGFINATSSQEEY 360
 QY 361 BIDPYEKFCFSVRKAYPRIRCTWIFPSQASFFCEQRLGEGDYSISKFCCHKPKGEIYIF 420
 Db 361 BIDPYEKFCFSVRKAYPRIRCTWIFPSQASFFCEQRLGEGDYSISKFCCHKPKGEIYIF 420

QY 421 AENDDAQFTKMTLNIRKKPQVLANASASQSSDGYPLPSWTWKCKSDKSPNCTEIP 480
 Db 421 AENDDAQFTKMTLNIRKKPQVLANASASQSSDGYPLPSWTWKCKSDKSPNCTEIP 480
 QY 481 EGVNKKANRKFVGQWVSSSTLNMSEAGKLLVKKCAVNSMGTSCTIFLNSPGPPFPFIQ 540
 Db 481 EGVNKKANRKFVGQWVSSSTLNMSEAGKLLVKKCAVNSMGTSCTIFLNSPGPPFPFIQ 540
 QY 541 DNISFYATIGLCPLPFIWLVILCHYKQKQFVRESQLOMTQVGTPLDNEFYVDFRDEY 600
 Db 541 DNISFYATIGLCPLPFIWLVILCHYKQKQFVRESQLOMTQVGTPLDNEFYVDFRDEY 600
 QY 601 DLKWEFPRENLEFGKVLGSGAFGRVNNATAYGISTGVSIOQAVKMLKEKADSCKEALM 660
 Db 601 DLKWEFPRENLEFGKVLGSGAFGRVNNATAYGISTGVSIOQAVKMLKEKADSCKEALM 660
 QY 661 SELKWMTHLGHDNINVLGACTLSGPVYLIFEYCCYGDLLNLRKREKPHRTWTBIFK 720
 Db 661 SELKWMTHLGHDNINVLGACTLSGPVYLIFEYCCYGDLLNLRKREKPHRTWTBIFK 720
 QY 721 EHNFSYPTFOAHNSNMPGSRVQLHPLDQLSGFGNGSIHSEDEIEYENQKRLAEEB 780
 Db 721 EHNFSYPTFOAHNSNMPGSRVQLHPLDQLSGFGNGSIHSEDEIEYENQKRLAEEB 780
 QY 781 EDNLVLTFFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVKVLCDFGLARDIL 840
 Db 781 EDNLVLTFFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVKVLCDFGLARDIL 840
 QY 841 SDSSYVVRGNARLPVKWMAPESLFEGIYTKSDVMSYGILLWEIFSLGVNPPYGPVDAN 900
 Db 841 SDSSYVVRGNARLPVKWMAPESLFEGIYTKSDVMSYGILLWEIFSLGVNPPYGPVDAN 900
 QY 901 FYKLQSGFKMQPPFYATGIVFMQSCWAFDSRKRSPFNLTSLFLGCQLAAEAEACIRT 960
 Db 901 FYKLQSGFKMQPPFYATGIVFMQSCWAFDSRKRSPFNLTSLFLGCQLAAEAEACIRT 960
 QY 961 SIHLPKQAAPQORGGRLAQSPOQVKIHRERS 992
 Db 961 SIHLPKQAAPQORGGRLAQSPOQVKIHRERS 992
 RESULT 5
 AAW19874
 ID AAW19874 standard; Protein; 992 AA.
 XX AAW19874;
 AC AC
 DT 25-MAR-2003 (updated)
 DT 19-AUG-1997 (first entry)
 XX
 DE Murine flk-2 receptor.
 XX
 KW Human; fetal liver kinase 2; flk2; receptor protein tyrosine kinase;
 KW pTK; liver; spleen; thymus; adult; brain; bone marrow; primitive portion;
 KW haematopoietic hierarchy; extracellular domain; soluble form; ligand;
 KW proliferation; differentiation; mammalian; haematopoietic stem cell;
 KW macrocytic anaemia; aplastic anaemia; cancer; radiation; mouse.
 XX
 OS Mus musculus.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..27
 FT /note= "Signal peptide"
 FT Domain 28..544
 FT /note= "Extracellular receptor domain"
 FT Region 545..564
 FT /note= "transmembrane region"
 FT Domain 565..992
 FT /note= "Intracellular catalytic domain"
 FT Active-site 618..623
 FT /note= "Catalytic sub-domain"
 FT Active-site 811..819

FT Active-site /note= "Catalytic sub-domain"
 FT 832..834
 FT Active-site /note= "Catalytic sub-domain"
 FT 857..862
 FT Active-site /note= "Catalytic sub-domain"
 FT 872..878
 FT Active-site /note= "Catalytic sub-domain"
 FT 709..785
 FT Peptide /note= "Signature sequence characteristic of flk-2"
 XX US5621090-A.
 PN 15-APR-1997.
 PD 26-JUN-1992; 92US-0906397.
 XX 26-JUN-1992; 92US-0906397.
 PR 02-APR-1991; 91US-0679666.
 PR 28-JUN-1991; 91US-0728913.
 PR 15-NOV-1991; 91US-0793065.
 PR 24-DEC-1991; 91US-0813593.
 XX (UYPR-) UNIV PRINCETON.
 PA Lemischka IR;
 PI WPI; 1997-235228/21.
 DR N-PSDB; AAT72118.
 XX Protein containing the extracellular domain of human flk-2 - used
 PT for identification of primitive haematopoietic cell proliferation
 PT and differentiation stimulatory ligands, e.g. for treating anaemia
 XX Disclosure; Fig 1A; 55pp; English.
 PS This sequence represents murine fetal liver kinase 2 (flk2). flk-2 is
 XX a receptor protein tyrosine kinase (PTK) and is important in transducing
 CC putative self-renewal signals from the environment. flk-2 is expressed
 CC in fetal liver, spleen, fetal thymus, adult brain and adult bone marrow,
 CC and it is thought that flk-2 is expressed in the entire primitive portion
 CC of the haematopoietic hierarchy. The invention concerns a recombinant
 CC nucleic acid, preferably mRNA, which encodes a protein containing only
 CC the extracellular domain of human flk-2 and lacking the flk-2 intra-
 CC cellular catalytic domain. The resultant protein represents a soluble
 CC form of flk-2 which is used to isolate specific ligands for flk-2. These
 CC ligands can be used to stimulate proliferation and/or differentiation of
 CC mammalian haematopoietic stem cells, in vivo or in vitro, e.g. for
 CC treatment of macrocytic or aplastic anaemia or bone marrow damage caused
 CC by cancer treatment or radiation.
 CC (Updated on 25-MAR-2003 to correct PF field.)
 XX Sequence 992 AA;
 SQ

Query Match 100.0%; Score 5264; DB 18; Length 992;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 992; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRALQSRDRRLLLVLSVMILETTNODLPVKVLISHENNGSSAGKPSYRVRGS 60
 DB 1 MRALQSRDRRLLLVLSVMILETTNODLPVKVLISHENNGSSAGKPSYRVRGS 60
 QY 61 PEDLOCTPRQSEGTVEAATVEAEGSITLQVQLATPGDLSCLVFKHSSLCQPHFD 120
 DB 61 PEDLOCTPRQSEGTVEAATVEAEGSITLQVQLATPGDLSCLVFKHSSLCQPHFD 120
 QY 121 LQNRGIVSMALNVTTQAGEYLLHIOSEANVTYVLTNNVDTQYVLRPFRQVQ 180
 DB 121 LQNRGIVSMALNVTTQAGEYLLHIOSEANVTYVLTNNVDTQYVLRPFRQVQ 180
 QY 181 DALLCISEGVPEPTVEWVLCSSHRECKEKGPAVRKEEVLHFGTDIRCCARNALGR 240
 DB 181 DALLCISEGVPEPTVEWVLCSSHRECKEKGPAVRKEEVLHFGTDIRCCARNALGR 240

QY 241 ECTKLTFTIDLNOAQPOSTLQPLFLKVGEPMLIRKCAIHVNHGFGLTWEDEKALBEGSYFE 300
 DB 241 ECTKLTFTIDLNOAQPOSTLQPLFLKVGEPMLIRKCAIHVNHGFGLTWEDEKALBEGSYFE 300
 QY 301 MSTYSTNRTMIRILLAFVSSVGRNDGYTTCSSSKHPSQSALVTILKGFNATSSQBEY 360
 DB 301 MSTYSTNRTMIRILLAFVSSVGRNDGYTTCSSSKHPSQSALVTILKGFNATSSQBEY 360
 QY 361 EIDPYEKFCFSVRKAYPRIRCTWIFSQASFPCEQGLEGDYGISKFCDHKNKPGEIFY 420
 DB 361 EIDPYEKFCFSVRKAYPRIRCTWIFSQASFPCEQGLEGDYGISKFCDHKNKPGEIFY 420
 QY 421 AENDDAQFTKMFNTINIRKPKQVLANASASQASCSDDGYPLPSWTWKCKSDKSPNCTBEIP 480
 DB 421 AENDDAQFTKMFNTINIRKPKQVLANASASQASCSDDGYPLPSWTWKCKSDKSPNCTBEIP 480
 QY 481 EGVNKKANRKFQGVQWSSSTLNNSEAGKLLVKCCAYNSMGTSCTIFLNSPGPPPIQ 540
 DB 481 EGVNKKANRKFQGVQWSSSTLNNSEAGKLLVKCCAYNSMGTSCTIFLNSPGPPPIQ 540
 QY 541 DNISFYATIGLCLPFIWLVILVLIHCHYKQFRYESQLQMTQVTPLDNEYFYVDYDEY 600
 DB 541 DNISFYATIGLCLPFIWLVILVLIHCHYKQFRYESQLQMTQVTPLDNEYFYVDYDEY 600
 QY 601 DLKWEFPRENLEBFGKVLGSGAFGRVNNATAYGISTGVSIOQAVKMLKEKADSCKEALM 660
 DB 601 DLKWEFPRENLEBFGKVLGSGAFGRVNNATAYGISTGVSIOQAVKMLKEKADSCKEALM 660
 QY 661 SELKMTHLGHNDINVLJGACTLSGPVYLIREYCCVGDLLNLYLSKREKPHRTWTIEFK 720
 DB 661 SELKMTHLGHNDINVLJGACTLSGPVYLIREYCCVGDLLNLYLSKREKPHRTWTIEFK 720
 QY 721 EHNFSYPTFOAHSNSSMPGSRVQLPDLQSLGFNGNSIHSEDEIEYENQKRLAEEEE 780
 DB 721 EHNFSYPTFOAHSNSSMPGSRVQLPDLQSLGFNGNSIHSEDEIEYENQKRLAEEEE 780
 QY 781 EDNLVLTFDLLCFAQVAKGMEFLFKSCVHRDLAARNVLVTHGKVKIKCDFGLARDIL 840
 DB 781 EDNLVLTFDLLCFAQVAKGMEFLFKSCVHRDLAARNVLVTHGKVKIKCDFGLARDIL 840
 QY 841 SDSSVYVRGNARLPVKWMAPESLFEGITYIKSDVMSYGILLWEIFSLGVNYPGIPVDAN 900
 DB 841 SDSSVYVRGNARLPVKWMAPESLFEGITYIKSDVMSYGILLWEIFSLGVNYPGIPVDAN 900
 QY 901 FYKLIQSGFKMEQPFYATIGIYFVMQSCWAFDSRKRKPSPNLTSFLGCQLAAEEACIRT 960
 DB 901 FYKLIQSGFKMEQPFYATIGIYFVMQSCWAFDSRKRKPSPNLTSFLGCQLAAEEACIRT 960
 QY 961 SIHLPKQAAPQQRGGLRAQSPQRVKIHRERS 992
 DB 961 SIHLPKQAAPQQRGGLRAQSPQRVKIHRERS 992

RESULT 6
 AAY08616
 ID AAY08616 standard; Protein; 992 AA.
 XX AC AAY08616;
 XX AC AAY08616;
 DT 05-AUG-1999 (first entry)
 XX Murine flk-2 protein.
 DE Murine flk-2; flk-1; cell isolation; fetal liver kinase; receptor;
 KW monoclonal; polyclonal; antibody; tyrosine kinase.
 OS Mus sp.
 XX US5912133-A.
 XX 15-JUN-1999.
 XX 10-FEB-1998; 98US-0021324.

XX 19-NOV-1992; 92US-0977451.
 PR 02-APR-1991; 91US-0679666.
 PR 28-JUN-1991; 91US-0728913.
 PR 15-NOV-1991; 91US-0793065.
 PR 24-DEC-1991; 91US-0813593.
 PR 26-JUN-1992; 92US-0906397.
 PR 12-NOV-1992; 92US-0975049.
 PR 30-APR-1993; 93US-0055269.
 PR 31-OCT-1994; 94US-0252498.
 PR 15-FEB-1996; 96US-0601891.
 XX (UYPR-) UNIV PRINCETON.
 XX Lemischka IR;
 XX WPI; 1999-357194/30.
 DR N-PSDB; AAX77514.
 XX Isolating hematopoietic cells expressing fetal liver kinase 1
 PT receptors
 XX Disclosure; Fig 1a; 59pp; English.
 XX This invention describes a novel method of isolating cells expressing
 CC fetal liver kinase 1 (flk-1) receptors on their surface and comprises
 CC binding the cells to a polyclonal or monoclonal antibody specific to
 CC the flk-1 receptor and isolating the cells that have bound to the
 CC antibody. The method can be used to isolate hematopoietic stem cells in
 CC any mammal but preferably a rat, mouse, rabbit or human. The proteins of
 CC the invention belong to the receptor protein family. This sequence
 CC represents the murine flk-2 protein which is used in the method of
 CC the invention.
 XX Sequence 992 AA;
 SQ

Query Match 100.0%; Score 5264; DB 20; Length 992;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 992; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MRALAQRDRRLLLVLSVLMILEVTNODLPVVKCVLISHENNGSSAGKPSYRMVRGS 60
 Db 1 MRALAQRDRRLLLVLSVLMILEVTNODLPVVKCVLISHENNGSSAGKPSYRMVRGS 60

Qy 61 PEDLOCTPRQSGFTVEAATVEAEGSITLQVQLATPGDLSCLVFKHSSLGCPHFD 120
 Db 61 PEDLOCTPRQSGFTVEAATVEAEGSITLQVQLATPGDLSCLVFKHSSLGCPHFD 120

Qy 121 LQNRGIVSMALNVETQAGEYLLHIQSERANYTLFTVNRDITOLYVLRPFRKMNQ 180
 Db 121 LQNRGIVSMALNVETQAGEYLLHIQSERANYTLFTVNRDITOLYVLRPFRKMNQ 180

Qy 181 DALLCISGVPEPTVEWVLCSSHRESCKEKGPAVVRKEEVLHFGTDIRCCARNALGR 240
 Db 181 DALLCISGVPEPTVEWVLCSSHRESCKEKGPAVVRKEEVLHFGTDIRCCARNALGR 240

Qy 241 ECTKLFTIDLNOAQPOSTLPQLKVGEPWIRCKAIHNVHGFGLTWELEDKALBERGSYFE 300
 Db 241 ECTKLFTIDLNOAQPOSTLPQLKVGEPWIRCKAIHNVHGFGLTWELEDKALBERGSYFE 300

Qy 301 MSTYSTNRMTIRILLAFVSSVGRNDTGYTSSSKHPSQSALVTILEKGFNATSSQBEY 360
 Db 301 MSTYSTNRMTIRILLAFVSSVGRNDTGYTSSSKHPSQSALVTILEKGFNATSSQBEY 360

Qy 361 EIDPYEKFCSVRKAYPRICTWIFSOASPCQORGLDGYISKFCDHKNKGEYIFY 420
 Db 361 EIDPYEKFCSVRKAYPRICTWIFSOASPCQORGLDGYISKFCDHKNKGEYIFY 420

Qy 421 AENDDAQFTKMTLNIRKKPQVLANASQASCSGSDGYPLFSWTWKKCSKSPNCTEIP 480
 Db 421 AENDDAQFTKMTLNIRKKPQVLANASQASCSGSDGYPLFSWTWKKCSKSPNCTEIP 480

Qy 481 EGVNKKANRKVFQGVSSSTLNMSSEAGKLLVKKCCAYNSMGTCSETIFLNSPGFPPIQ 540

Db 481 EGVNKKANRKVFQGVSSSTLNMSSEAGKLLVKKCCAYNSMGTCSETIFLNSPGFPPIQ 540
 Qy 541 DNISFYATIGCLPPIVVLIVLICHYKKQPRYESQLOMIQVTPLDNNEYFYVDFRDY 600
 Db 541 DNISFYATIGCLPPIVVLIVLICHYKKQPRYESQLOMIQVTPLDNNEYFYVDFRDY 600
 Qy 601 DLKWEFPRENLEFGKVLGSGAGFRVMNATAYGISTGYSIQVAVKMLKEKADSCKEALM 660
 Db 601 DLKWEFPRENLEFGKVLGSGAGFRVMNATAYGISTGYSIQVAVKMLKEKADSCKEALM 660
 Qy 661 SELKXMTLGHHDNIVNLGACTLSGPVYLIFEYCCYDGLLNLYRSKRKPHRTWTEIFK 720
 Db 661 SELKXMTLGHHDNIVNLGACTLSGPVYLIFEYCCYDGLLNLYRSKRKPHRTWTEIFK 720
 Qy 721 EHNFSSTYPTFOAHNSNMPGSRVQLHPPLDQLSGFNNGSIHSEDEIEYENQKRLAESEE 780
 Db 721 EHNFSSTYPTFOAHNSNMPGSRVQLHPPLDQLSGFNNGSIHSEDEIEYENQKRLAESEE 780
 Qy 781 EDLNVLTPEDLICFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVKICDFGLARDIL 840
 Db 781 EDLNVLTPEDLICFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVKICDFGLARDIL 840
 Qy 841 SDSSVVRGNARLPVKWMAPESLPEGIYTIKSDVMSYGILLWEIFSLGWNYPGIPVDAN 900
 Db 841 SDSSVVRGNARLPVKWMAPESLPEGIYTIKSDVMSYGILLWEIFSLGWNYPGIPVDAN 900
 Qy 901 FYKLIQSFGKMEQFPYATEGTYFYVQWSCWAFDSRKRPSFPNLTSLFGQLAEAEACIRT 960
 Db 901 FYKLIQSFGKMEQFPYATEGTYFYVQWSCWAFDSRKRPSFPNLTSLFGQLAEAEACIRT 960
 Qy 961 SIHLPKQAAPOQRGLRAQSPQRQVKIHRERS 992
 Db 961 SIHLPKQAAPOQRGLRAQSPQRQVKIHRERS 992

RESULT 7
 ABG70915
 ID ABG70915 standard; Protein; 992 AA.
 XX AC ABG70915;
 XX DT 09-DEC-2002 (first entry)
 XX Mouse receptor protein tyrosine kinase, FLK-2.
 DE Mouse; FLK-2; foetal liver kinase; antianaemic; enzyme;
 KW opthalmological; receptor protein tyrosine kinase; aplastic anaemia;
 KW primitive haematopoietic cell; stem cell; macrocytic anaemia;
 KW bone marrow damage; cancer chemotherapy.
 XX Mus sp.
 XX Key Location/Qualifiers
 FH Peptide 1..27
 FT Protein /label= Signal_peptide
 FT 28..992 /label= Mature_FLK_2
 XX US2002119545-A1.
 XX 29-AUG-2002.
 XX 01-JUN-2001; 2001US-0872136.
 XX 19-NOV-1992; 92US-0977451.
 PR 30-APR-1993; 93US-0055269.
 PR 31-OCT-1994; 94US-0252498.
 PR 15-FEB-1996; 96US-0601891.
 PR 10-FEB-1998; 98US-0021324.
 PR 10-DEC-1998; 98US-0208786.
 PR 02-APR-1991; 91US-0679666.
 PR 28-JUN-1991; 91US-0728913.

PR 15-NOV-1991; 91US-0793065.
PR 24-DEC-1991; 91US-0813593.
XX (LEMI/) LEMISCHKA I R.
XX Lemischka IR;
XX WPI; 2002-731356/79.
DR N-PSDB; ABS55044.
XX
XX New protein tyrosine kinase expressed in primitive hematopoietic cells
PT (HC) and not expressed in mature HC, and ligands for the protein,
PT useful for stimulating proliferation of primitive hematopoietic stem
PT cells -
XX
XX Claim 37; Fig 1a; 64pp; English.
XX
XX The invention relates to a receptor protein tyrosine kinase (pTK)
CC expressed in primitive hematopoietic cells and not expressed in mature
CC hematopoietic cells, and named FLK-1 and -2 (foetal liver kinase).
CC Also included are the nucleic acids encoding the FLK proteins, FLK
CC expression vectors, a ligand that binds to human or murine FLK-2 or FLK-1
CC and stimulates the proliferation and/or differentiation of the primitive
CC hematopoietic cells and a murine cell line 2018 having American Type
CC Culture Collection (ATCC) accession number ATCC CRL 10907. The ligands
CC are useful for stimulating the proliferation and/or differentiation of
CC primitive mammalian hematopoietic stem cells. The receptor pTK molecules
CC are useful for stimulating the self renewal of the totipotent
CC hematopoietic stem cell and to stimulate the development of all cells of
CC the hematopoietic system both in vitro and in vivo. The ligands for the
CC receptors act as hematopoietic growth factors. The ligands are useful in
CC treating humans whose primitive stem cells do not sufficiently undergo
CC self-renewal e.g. macrocytic and aplastic anaemia and bone marrow damage
CC resulting from cancer chemotherapy and radiation. The present
CC sequence represents mouse FLK-2.
XX
XX Sequence 992 AA;

Query Match 100.0%; Score 5264; DB 23; Length 992;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 992; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRALQSRDRLLLLVLSVILETVNODLPVKVILSHENNGSSAGKPSYRMVRS 60
DB 1 MRALQSRDRLLLLVLSVILETVNODLPVKVILSHENNGSSAGKPSYRMVRS 60
QY 61 PEDLOCTPRQSEGTVEAATVEAEGSITLQVQLATPGDLSCLVFKHSGICOPHPD 120
DB 61 PEDLOCTPRQSEGTVEAATVEAEGSITLQVQLATPGDLSCLVFKHSGICOPHPD 120
QY 121 LQNRGIVSMALNVETQAGEYLLHIQSERANYTLFTVNVVDTQYLVLRFPYRQWENQ 180
DB 121 LQNRGIVSMALNVETQAGEYLLHIQSERANYTLFTVNVVDTQYLVLRFPYRQWENQ 180
QY 181 DALLCISEGVPEPTVEWVLCSSHRECKEGBPAVVRKEKVLHELFGTDIRCARNALGR 240
DB 181 DALLCISEGVPEPTVEWVLCSSHRECKEGBPAVVRKEKVLHELFGTDIRCARNALGR 240
QY 241 ECTKLTIDLNQAPOSTLPQLFKVGEPLWIRCKAIHNVNHHGFLTWELDKALEGSYPE 300
DB 241 ECTKLTIDLNQAPOSTLPQLFKVGEPLWIRCKAIHNVNHHGFLTWELDKALEGSYPE 300
QY 301 MSTYTNTRMTIRILLAFVSSVGRNDTGYTCSSSHPSQSALVTILEKGFINATSSQERY 360
DB 301 MSTYTNTRMTIRILLAFVSSVGRNDTGYTCSSSHPSQSALVTILEKGFINATSSQERY 360
QY 361 EIDPYEKFCSVRKAYPIRCTWIFSQASFPCEQRGLEDGYISIKFCDHKNKPGYIIFY 420
DB 361 EIDPYEKFCSVRKAYPIRCTWIFSQASFPCEQRGLEDGYISIKFCDHKNKPGYIIFY 420
QY 421 AENDDAQFTKMTLIRKKPOVLANASASQASCSDDGYPLPSWTWKCSDKSPNCTEIP 480
DB 421 AENDDAQFTKMTLIRKKPOVLANASASQASCSDDGYPLPSWTWKCSDKSPNCTEIP 480

QY 481 EGVWNNKANRKYFGQWSSSTLNMSGAKGLLVKCCAYNSMGTSCETIFLNSPGPPFFIQ 540
DB 481 EGVWNNKANRKYFGQWSSSTLNMSGAKGLLVKCCAYNSMGTSCETIFLNSPGPPFFIQ 540
QY 541 DNISFYATIGLCLPIVVLVLI CHYKQKQFYESOLOMIQVTPGLDNEFYVYDFRDY 600
DB 541 DNISFYATIGLCLPIVVLVLI CHYKQKQFYESOLOMIQVTPGLDNEFYVYDFRDY 600
QY 601 DLKWEFFPRENLEFGKVLGSGAFGRVNNATAYGISKTGVSIOAVXMKLKEKADSCKEALM 660
DB 601 DLKWEFFPRENLEFGKVLGSGAFGRVNNATAYGISKTGVSIOAVXMKLKEKADSCKEALM 660
QY 661 SELKQWTHLGHNDTNVNLGACTLSPVYLIPEYCCYGDLLNLYLRSKRKFHRTWTEIFK 720
DB 661 SELKQWTHLGHNDTNVNLGACTLSPVYLIPEYCCYGDLLNLYLRSKRKFHRTWTEIFK 720
QY 721 EHNFSYPTFOAHSSNMPGSGREVOLHPDOLSGFNGNSIHSEDEIEYENOKRLAESEE 780
DB 721 EHNFSYPTFOAHSSNMPGSGREVOLHPDOLSGFNGNSIHSEDEIEYENOKRLAESEE 780
QY 781 EDNLVLTTFEDLLCFAYQVAKGMFELEFKSCVHRDLAARNVLTGHGVKVICDFGLARDIL 840
DB 781 EDNLVLTTFEDLLCFAYQVAKGMFELEFKSCVHRDLAARNVLTGHGVKVICDFGLARDIL 840
QY 841 SDSSYVVRGNARLPVKWMAPELSLFGIYTIKSDVMSYGILLWEIFSLGVNYPGIPVDAN 900
DB 841 SDSSYVVRGNARLPVKWMAPELSLFGIYTIKSDVMSYGILLWEIFSLGVNYPGIPVDAN 900
QY 901 FYKLQSGFKMOPVYATEGIFVWQSWAFDSRKRSPFNLTSLGCOLAEAEACIRT 960
DB 901 FYKLQSGFKMOPVYATEGIFVWQSWAFDSRKRSPFNLTSLGCOLAEAEACIRT 960
QY 961 SIHLPKQAAPQORGGLRAQSPORQVKIHRERS 992
DB 961 SIHLPKQAAPQORGGLRAQSPORQVKIHRERS 992

RESULT 8
AAE25818
ID AAE25818 standard; Protein; 992 AA.
XX AAE25818;
XX 15-NOV-2002 (first entry)
XX Murine receptor protein tyrosine kinase, flk-2.
DE Murine; receptor protein tyrosine kinase; pTK; haematopoietic cell;
KW growth factor; therapeutic; macrocytic anaemia; aplastic anaemia;
KW bone marrow damage; cancer; chemotherapy; radiation; osteopathic;
KW flk-2 protein.
XX Mus sp.
XX Key Location/Qualifiers
FH Peptide 1..27
FT /label= Signal-peptide
FT Protein 28..992
FT /note= "Murine mature flk-2 protein"
FT Domain 1..517
FT /note= "Extracellular domain"
FT Domain 518..537
FT /note= "Transmembrane domain"
FT Domain 538..966
FT /note= "Intracellular domain"
XX US2002072077-A1.
XX 13-JUN-2002.
XX 31-JUL-2001; 2001US-0919408.
XX

PR 19-NOV-1992; 92US-0977451.
 PR 30-APR-1993; 93US-0055269.
 PR 31-OCT-1994; 94US-0252498.
 PR 15-FEB-1996; 96US-0601891.
 PR 10-FEB-1998; 98US-0021324.
 PR 10-DEC-1998; 98US-0208786.
 PR 02-APR-1991; 91US-0679666.
 PR 28-JUN-1991; 91US-0728913.
 PR 15-NOV-1991; 91US-0793065.
 PR 24-DEC-1991; 91US-0813593.
 XX (LEMI/) LEMISCHKA I R.
 XX Lemischka IR;
 DR WPI; 2002-607237/65.
 DR N-PSDB; AAD42483.
 XX
 PT New protein tyrosine kinase expressed in primitive hematopoietic cells
 PT (HC) and not expressed in mature HC, and ligands for the protein, for
 PT stimulating proliferation of primitive hematopoietic stem cells -
 XX
 PS Claim 37; Page 19-21; 64pp; English.
 XX
 CC The present invention relates to receptor protein tyrosine kinase (PTK)
 CC expressed in primitive hematopoietic cells and not expressed in mature
 CC hematopoietic cells, polynucleotides encoding such proteins and ligands
 CC for the protein. Ligands which bind to PTK of the invention are useful
 CC for stimulating the proliferation and/or differentiation of primitive
 CC mammalian hematopoietic stem cells. The receptor PTK molecules are
 CC useful for stimulating the self renewal of the totipotent hematopoietic
 CC stem cell and to stimulate the development of all cells of hematopoietic
 CC system both in vitro and in vivo. The ligands for the receptors act as
 CC hematopoietic growth factors. The ability of the ligands to stimulate
 CC proliferation of stem cells both in vitro and in vivo has important
 CC therapeutic applications such as treating humans whose primitive stem
 CC cells do not sufficiently undergo self-renewal. It is also useful in
 CC conditions that occur when defects in hematopoietic stem cells or their
 CC related growth factors depress the number of white blood cells such as
 CC macrocytic and aplastic anaemia or bone marrow damage resulting from
 CC cancer chemotherapy and radiation. The present sequence is murine
 CC receptor PTK, flk-2.
 XX
 SQ Sequence 992 AA;

Query Match 100.08; Score 5264; DB 23; Length 992;
 Best Local Similarity 100.08; Pred. No. 0;
 Matches 992; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRALAQRSDRLLLVLSVMILETVTNQDLPVVKCVLISHENNGSSAGKPSYRMVGRS 60
 DB 1 MRALAQRSDRLLLVLSVMILETVTNQDLPVVKCVLISHENNGSSAGKPSYRMVGRS 60
 QY 61 PEDLOCTPRROSEGTVEAATVEAEGSITLQVQLATPGDLSCILWFKHSGLGCPHFD 120
 DB 61 PEDLOCTPRROSEGTVEAATVEAEGSITLQVQLATPGDLSCILWFKHSGLGCPHFD 120
 QY 121 LQNRGIVSMALNTVETQAGYILHIQSERANYTVLFTVNRDQYVLRYPYFRKMNQ 180
 DB 121 LQNRGIVSMALNTVETQAGYILHIQSERANYTVLFTVNRDQYVLRYPYFRKMNQ 180
 QY 181 DALLCISGEVPEPTVWVLCSSHRESCKECPAVVRKEEVLHFLFGTDIRCCARNALGR 240
 DB 181 DALLCISGEVPEPTVWVLCSSHRESCKECPAVVRKEEVLHFLFGTDIRCCARNALGR 240
 QY 241 ECTKLFTIDLNQAPQSTLPQLFLKVGEPFLIRCAIHNHGFGLTWLEDKALEEGSVFE 300
 DB 241 ECTKLFTIDLNQAPQSTLPQLFLKVGEPFLIRCAIHNHGFGLTWLEDKALEEGSVFE 300
 QY 301 MSTYSTNRTMIRILLAFVSSVGRNDTGYTCSSSKHPSQSALVTILEKGFINATSSQEEY 360
 DB 301 MSTYSTNRTMIRILLAFVSSVGRNDTGYTCSSSKHPSQSALVTILEKGFINATSSQEEY 360

QY 361 EIDPYEKFCFSVRFKAYPRIRCTWIFSOASPPCEORGLEDGYSISKCDHKVKGEYIFY 420
 DB 361 EIDPYEKFCFSVRFKAYPRIRCTWIFSOASPPCEORGLEDGYSISKCDHKVKGEYIFY 420
 QY 421 AENDDAQFTKMFTLNIRKKQVLANASASQASCSDDGYPLPSWTWKCKSDKSPNCTEIP 480
 DB 421 AENDDAQFTKMFTLNIRKKQVLANASASQASCSDDGYPLPSWTWKCKSDKSPNCTEIP 480
 QY 481 EGVNKKANKRVFGQWVSSSTLNSBAGKLLVKCCAYNSMGTSCTETFLNSPGFPFPIQ 540
 DB 481 EGVNKKANKRVFGQWVSSSTLNSBAGKLLVKCCAYNSMGTSCTETFLNSPGFPFPIQ 540
 QY 541 DNISFYATIGLCPLPFIWLVILCHIKYKQFRIESQLOMIQVGTPLDNEYFYVDFRDEY 600
 DB 541 DNISFYATIGLCPLPFIWLVILCHIKYKQFRIESQLOMIQVGTPLDNEYFYVDFRDEY 600
 QY 601 DLKWEFPRENLEFGKVLGSGAFGRVNNATAYGISKTVGSIQAVKMLKADSCKEKALM 660
 DB 601 DLKWEFPRENLEFGKVLGSGAFGRVNNATAYGISKTVGSIQAVKMLKADSCKEKALM 660
 QY 661 SELKXMTLGHNDINVLGACTLSGPVYLIFEYCCYGDILLNLRKREKFRHTWTWEIFK 720
 DB 661 SELKXMTLGHNDINVLGACTLSGPVYLIFEYCCYGDILLNLRKREKFRHTWTWEIFK 720
 QY 721 EHNFSYPTFOAHSSNMFSGSREVQLHPPDLQSLGFGNSIHSDEIEYENQKRLAESEE 780
 DB 721 EHNFSYPTFOAHSSNMFSGSREVQLHPPDLQSLGFGNSIHSDEIEYENQKRLAESEE 780
 QY 781 EDLNVLTPEDLACFAVQVAKGMEFEFKSCVHRDLAARNVLVTHGKVKICDFGLARDIL 840
 DB 781 EDLNVLTPEDLACFAVQVAKGMEFEFKSCVHRDLAARNVLVTHGKVKICDFGLARDIL 840
 QY 841 SDSSYVVRGNARLPVKWAPESLFGIYTIKSDVMSYGILLWEIFSLGVNYPYGPVPDAN 900
 DB 841 SDSSYVVRGNARLPVKWAPESLFGIYTIKSDVMSYGILLWEIFSLGVNYPYGPVPDAN 900
 QY 901 FYKLQSGFKMEQPPYATEGIVFMQSCWAFDSRKRPSFPNLTSLGCOLAEAEACIRT 960
 DB 901 FYKLQSGFKMEQPPYATEGIVFMQSCWAFDSRKRPSFPNLTSLGCOLAEAEACIRT 960
 QY 961 SIHLPKQAPQORGLRAQSPQORVKIHRERS 992
 DB 961 SIHLPKQAPQORGLRAQSPQORVKIHRERS 992

RESULT 9
 AAR44994
 ID AAR44994 standard; Protein; 992 AA.
 XX AAR44994;
 AC
 XX
 DT 25-MAR-2003 (updated)
 DT 27-JUN-1994 (first entry)
 DE
 XX
 XX Murine flk-2 receptor protein tyrosine kinase.
 KW Receptor protein tyrosine kinase; pTK family; foetal liver kinase;
 KW primitive; totipotent; haematopoietic cell; stem cell; proliferation;
 KW mflk; stromal cell.
 XX
 OS Mus musculus.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..27
 FT /label= signal_peptide
 FT /note= "hydrophobic leader"
 FT Protein 28..992
 FT /label= flk-2
 FT /note= "mature protein"
 FT Domain 28..544
 FT /label= extracellular_domain
 FT Region 545..564
 FT /label= transmembrane_region

FT Region /note= "Extracellular receptor domain"
 FT 545..564
 FT Domain /note= "Transmembrane region"
 FT 565..992
 FT Domain /note= "Intracellular catalytic domain"
 FT 545..564
 FT Domain /note= "Catalytic subdomain #1"
 FT 618..623
 FT Domain /note= "Catalytic subdomain #2"
 FT 811..819
 FT Domain /note= "Catalytic subdomain #3"
 FT 832..834
 FT Domain /note= "Catalytic subdomain #4"
 FT 857..862
 FT Domain /note= "Catalytic subdomain #5"
 FT 872..878
 FT Domain /note= "Catalytic subdomain #6"
 FT 709..785
 FT Region /note= "Signature sequence characteristic of flk-2"
 XX WO9300349-A1.
 XX 07-JAN-1993.
 XX 26-JUN-1992; 92WO-US05401.
 XX 28-JUN-1991; 91US-0728913.
 XX 15-NOV-1991; 91US-0793065.
 XX 24-DEC-1991; 91US-0813593.
 XX 02-APR-1992; 92WO-US02750.
 XX (UYPR-) UNIV PRINCETON.
 XX Lemischka IR;
 XX WPI; 1993-036323/04.
 XX N-PSDB; AAQ45247.
 XX Nucleic acid encoding receptor protein tyrosine kinase - allows
 XX development of ligands to stimulate proliferation and/or
 XX differentiation of mammalian haematopoietic stem cells
 XX Claim 5; Fig 1a; 78pp; English.
 XX This sequence represents a murine receptor protein tyrosine kinase
 XX which belongs to a new functional class of protein tyrosine kinases
 XX (PTKs). PTKs in this class are expressed in primitive mammalian
 XX hematopoietic (pHC) cells but not in mature hematopoietic cells (mHC).
 XX This protein is an example of a receptor ptk and is called fetal liver
 XX kinase 2 (flk-2). flk-2 is expressed in fetal liver, spleen and
 XX thymus, and adult brain and marrow. Expression of flk-2 mRNA occurs
 XX in the most primitive thymocyte subset, which is believed to be
 XX uncommitted. Therefore, thymocytes expressing flk-2 may be multi-
 XX potential. flk-2 is the first receptor tyrosine kinase known to be
 XX expressed in the T-lymphoid lineage.
 XX (Updated on 25-MAR-2003 to correct PN field.)
 XX Sequence 992 AA;
 XX
 Query Match 99.8%; Score 5256; DB 14; Length 992;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 991; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MRALAQSRRLLLVLSWILEVTNQLPVIKCVLISHENNGSAGRPSSYRMVGRS 60
 DB 1 MRALAQSRRLLLVLSWILEVTNQLPVIKCVLISHENNGSAGRPSSYRMVGRS 60
 QY 61 PEDLOCTPRQSGTVEATVEASGSGTTLQVLAATPGDLSCLWVFKHSSIGCCPHFD 120
 DB 61 PEDLOCTPRQSGTVEATVEASGSGTTLQVLAATPGDLSCLWVFKHSSIGCCPHFD 120
 QY 121 LQNRGIVSMALINVTQAGEYLLHIQSERANYTVLFTVNRVDTQLYVLRPPYFRKWNQ 180

Db 121 LQNRGIVSMALINVTQAGEYLLHIQSERANYTVLFTVNRVDTQLYVLRPPYFRKWNQ 180
 QY 181 DALLCISGVPEPTVWVLCSSHRESCKEKGPAVVRKEKVLHFLGTDIRCARNALGR 240
 Db 181 DALLCISGVPEPTVWVLCSSHRESCKEKGPAVVRKEKVLHFLGTDIRCARNALGR 240
 QY 241 ECTKLFTIDLNAQPOSTLPQLFLKVGEPILWIRCKAIHVNHGFLTWELDKALEESYFE 300
 Db 241 ECTKLFTIDLNAQPOSTLPQLFLKVGEPILWIRCKAIHVNHGFLTWELDKALEESYFE 300
 QY 301 MSTYSTNRTMIRILLAFVSSVGRNDTGYTCCSSKHPSQSALVTILEKGFINATSQEY 360
 Db 301 MSTYSTNRTMIRILLAFVSSVGRNDTGYTCCSSKHPSQSALVTILEKGFINATSQEY 360
 QY 361 BIDPYEKCFVSVRKAYPRICTWIFSOASPCGEORGLDGYISIKFCHDKNKPGEYIPY 420
 Db 361 BIDPYEKCFVSVRKAYPRICTWIFSOASPCGEORGLDGYISIKFCHDKNKPGEYIPY 420
 QY 421 AENDDAQPTKMTLNIRKKPOVLANASASQASCSGSDGYPLPSWTWKKCSDKSNCTEIP 480
 Db 421 AENDDAQPTKMTLNIRKKPOVLANASASQASCSGSDGYPLPSWTWKKCSDKSNCTEIP 480
 QY 481 EGVNKKANRKFVQGVSSSTLNMSKAGKLLVKCCAYNSMGTSCTETIFLNSPGPFPTQ 540
 Db 481 EGVNKKANRKFVQGVSSSTLNMSKAGKLLVKCCAYNSMGTSCTETIFLNSPGPFPTQ 540
 QY 541 DNISFYATIGLCLPFIIVLVLI CHYKKQPRYSQLOMIQVTGPDNEYFYVDFRDY 600
 Db 541 DNISFYATIGLCLPFIIVLVLI CHYKKQPRYSQLOMIQVTGPDNEYFYVDFRDY 600
 QY 601 DLKWEPPRENLEFGKVLGSGAFGRVMNATAYGISTGVSIOVAVKMLKEKADSCKEALM 660
 Db 601 DLKWEPPRENLEFGKVLGSGAFGRVMNATAYGISTGVSIOVAVKMLKEKADSCKEALM 660
 QY 661 SELKMWTHLGHNDINVLGACTLSGPVYLIFCYCCYGDLLNLYRSKREKTHRTWTEIFK 720
 Db 661 SELKMWTHLGHNDINVLGACTLSGPVYLIFCYCCYGDLLNLYRSKREKTHRTWTEIFK 720
 QY 721 EHNPSYPTFOAHNSNMPGSRVOLHPPLDQLSGFNGNSIHSDEIEYENQKRLAEE 780
 Db 721 EHNPSYPTFOAHNSNMPGSRVOLHPPLDQLSGFNGNSIHSDEIEYENQKRLAEE 780
 QY 781 EDNLNLTFFEDLLCFAYQVAKMEFLPKSCVHRDLAARNVLVTHGKVVKICDFGLARDIL 840
 Db 781 EDNLNLTFFEDLLCFAYQVAKMEFLPKSCVHRDLAARNVLVTHGKVVKICDFGLARDIL 840
 QY 841 SDSYVVRGNARLPVKWMAPESEFEGYITTKSDVWSYGILLWEIFSLGVNPPYGPV 900
 Db 841 SDSYVVRGNARLPVKWMAPESEFEGYITTKSDVWSYGILLWEIFSLGVNPPYGPV 900
 QY 901 FYKLIQSGFQWQPFYVATEGIYFVMQSCWAFDSRKRPSPFNLTSLFLGCQLAEAEACIRT 960
 Db 901 FYKLIQSGFQWQPFYVATEGIYFVMQSCWAFDSRKRPSPFNLTSLFLGCQLAEAEACIRT 960
 QY 961 SIHLPKQAAPQORGLRAQSPQVKIHRERS 992
 Db 961 SIHLPKQAAPQORGLRAQSPQVKIHRERS 992
 RESULT 11
 AAR28038
 ID AAR28038 standard; Protein; 992 AA.
 XX AAR28038;
 XX AC
 XX XX
 DT 25-MAR-2003 (updated)
 DT 15-MAR-1993 (first entry)
 XX XX
 DE Murine flk-2.
 XX
 KW Thymidine kinase; TK; haematopoietic; stem cells; proliferation;
 KW differentiation; progenitor cells; foetal liver kinase.
 XX

OS	Mus musculus.
XX	
XX	Location/Qualifiers
PH	Key
FT	Peptide
FT	1..27
FT	/note= "hydrophobic leader"
FT	28...544
FT	/note= "extracellular receptor domain"
FT	545..564
FT	/note= "transmembrane region"
FT	565..992
FT	/note= "intracellular catalytic domain"
FT	709..785
FT	/note= "signature sequence of flk-2"
XX	WO9217486-A1.
XX	
PN	15-OCT-1992.
XX	
PD	
XX	
PF	02-APR-1992; 92WO-US02750.
XX	
XX	02-APR-1991; 91US-0679666.
PR	28-JUN-1991; 91US-0728913.
PR	15-NOV-1991; 91US-0793065.
PR	24-DEC-1991; 91US-0813593.
XX	
PA	(UYPR-) UNIV PRINCETON.
XX	
PI	Lemischka IR;
XX	
DR	WPI; 1992-366185/44.
DR	N-PSDB; AAQ29954.
XX	
PT	Stimulating proliferation and/or differentiation of primitive mammalian haematopoietic stem cells - using ligand that binds thymidine kinase and flk-1 and flk-2
PT	
PT	
XX	
PS	Claim 37; Fig 1a; 94pp; English.
CC	The murine foetal liver kinase (flk) -2 sequence was deduced from that of the flk-2 clone obtd. by standard PCR techniques from stem-cell receptor-contg. tissue cDNA libraries. Suitable tissues include foetal liver, spleen or thymus cells or adult marrow or brain cells. The PCR primers used are based on known sections of the flk-2 gene. The murine flk-2 clone may be used in a vector to transform haematopoietic cells. The thymidine kinase encoded by flk-2 is expressed in primitive but not mature haematopoietic cells. Ligand binding to the TK may be prepd. which can stimulate proliferation and/or differentiation of primitive haematopoietic cells in vivo. The ligands can stimulate the proliferation of additional primitive stem cells, differentiation into more mature progenitor cells, or both. See also AAR28039-41.
CC	(Updated on 25-MAR-2003 to correct PN field.)
XX	
SQ	Sequence 992 AA;
	Query Match 99.5%; Score 5239; DB 13; Length 992;
	Best Local Similarity 99.7%; Pred. No. 0;
	Matches 989; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Oy	1 MRALAQRSDRLLLLVLSVMILETVTQNQLPVIKCVLISHENNGSSAGKSPSSYRMVRGS 60
Db	
	1 MRALAQRSDRLLLLVLSVMILETVTQNQLPVIKCVLISHENNGSSAGKSPSSYRMVRGS 60
Oy	61 PEDLCQTPRROSEGVTVAAATVEAEGSITLQVLATPGDLSCLVWFKHSLGCGPHFD 120
Db	
	61 PEDLCQTPRROSEGVTVAAATVEAEGSITLQVLATPGDLSCLVWFKHSLGCGPHFD 120
Oy	121 LQNRGIYSMAILNTQTQAGYLLIHIOSE RANYTVLTWNVRDTQLYLRRPYFRKVENQ 180
Db	
	121 LQNRGIYSMAILNTQTQAGYLLIHIOSE RANYTVLTWNVRDTQLYLRRPYFRKVENQ 180
Oy	181 DALLCISIEGVPEPTVFWVWLCSHRESCKE GPA VVRKEEKVLHFPGTDIRCCARNALGR 240

XX PN WO9527062-A1.
 XX PD 12-OCT-1995.
 XX PF 23-MAR-1995; 95WO-US03718.
 XX PR 04-APR-1994; 94US-0222299.
 XX (GETH) GENENTECH INC.
 XX PI Bennett BD, Broz SD, Matthews W, Zeigler FC;
 XX DR WPI; 1995-358636/46.
 XX DR N-PSDB; AAT00801.
 XX PT Agonist antibody which binds to flk2-flt3 tyrosine kinase receptor
 PT - enhances proliferation of haematopoietic stem cells, in the
 PT treatment of hypoplasia, anaemia, etc.
 XX PS Example 1; Page 38-40; 59pp; English.
 XX CC Murine stem cell tyrosine kinase receptor flk2 (also called flt3)
 CC (AAR91868) is encoded DNA (AAR0801) obt'd. by RT-PCR amplification of
 CC RNA isolated from mid-gestation mouse foetal livers. An flk2/flt3
 CC extracellular domain-IgG1 FC fusion protein was used to raise
 CC agonist antibodies able to bind to, and activate, flk2/flt3.
 CC Such antibodies can enhance repopulation of mature blood cell
 CC lineages following chemotherapy, radiotherapy, or bone marrow
 CC transplantation.
 XX SQ Sequence 1000 AA;

Query Match 96.9%; Score 5102; DB 16; Length 1000;
 Best Local Similarity 96.7%; Pred. No. 0;
 Matches 968; Conservative 6; Mismatches 17; Indels 10; Gaps 2;

QY 1 MRALAQRDRRLLLVLSVMILETVNODLPVKVLISHENNGSSAGKPSYRVRGS 60
 DB 1 MRALAQRDRRLLLVLSVMILETVNODLPVKVLISHENNGSSAGKPSYRVRGS 60
 QY 61 PEDLOCTPRQSGTYVEATVEASGSIITLQVQLATPGDLSCLWVFKHSLGCGPHFD 120
 DB 61 PEDLOCTPRQSGTYVEATVEASGSIITLQVQLATPGDLSCLWVFKHSLGCGPHFD 120
 QY 121 LQNRGIVSMALLNVTTQAGEYLLHIQSEAAVTVLFTVNVRTQLYVLRPFYRKMENQ 180
 DB 121 LQNRGIVSMALLNVTTQAGEYLLHIQSEAAVTVLFTVNVRTQLYVLRPFYRKMENQ 180
 QY 181 DALLCISEGVPEPTVEWVLCSSHRESCKEGPAVVRKEEVLHFGTDIRCCARNALGR 240
 DB 181 DALLCISEGVPEPTVEWVLCSSHRESCKEGPAVVRKEEVLHFGTDIRCCARNALGR 240
 QY 241 ECTKLTIDLNOAQSTLPQLFKVGEPLWIRCKAIHNVHFGLTWELEDKALEEGSYFE 300
 DB 241 ESTKLTIDLNOAQSTLPQLFKVGEPLWIRCKAIHNVHFGLTWELEDKALEEGSYFE 300
 QY 301 MSTYSTNRTMIRILLAFVSSVGRNDTGYTTCSSSKHPSQSALVTILEKGFINATSSQEEY 360
 DB 301 MSTYSTNRTMIRILLAFVSSVGRNDTGYTTCSSSKHPSQSALVTILEKGFINATSSQEEY 360
 QY 361 EIDPYEKFCSVRFKAYPRIRCTWIFSQASFPCEQRLGDEGYSISKPCDHKNKPGYIY 420
 DB 361 EIDPYEKFCSVRFKAYPRIRCTWIFSQASFPCEQRLGDEGYSISKPCDHKNKPGYIY 420
 QY 421 AENDDAQFTKMTNLIRKPKQVLANASASQASCSGYPPLPSMTWKKCKSPNCTEEIP 480
 DB 421 AENDDAQFTKMTNLIRKPKQVLANASASQASCSGYPPLPSMTWKKCKSPNCTEEIP 480
 QY 481 EGVWNKANKRVFGQWVSSSTLNMSBAGLLVKCCAYNSMGTSCTIFLNSGPPPPFTQ 540
 DB 481 EGVWNKANKRVFGQWVSSSTLNMSBAGLLVKCCAYNSMGTSCTIFLNSGPPPPFTQ 540

QY 541 DNISFYATIGLCPLPFIIVLVILVILCHYKQPRYESQLQMIQVGTQPLDNEYFYVDFRDY 600
 DB 541 DNISFYATIGLCPLPFIIVLVILVILCHYKQPRYESQLQMIQVGTQPLDNEYFYVDFRDY 600
 QY 601 DLKWEFPRENLEFGKVLGSGAFGRVMNATAYGISKTGVSQVAVKMLKEKADSCKEALM 660
 DB 601 DLKWEFPRENLEFGKVLGSGAFGRVMNATAYGISKTGVSQVAVKMLKEKADSCKEALM 660
 QY 661 SELKMMTHLGHNDINVLGACTLSGPVYLIFCYCCYGDLLNLYLRSKREKPHRTWTEIFK 720
 DB 661 SELKMMTHLGHNDINVLGACTLSGPVYLIFCYCCYGDLLNLYLRSKREKPHRTWTEIFK 720
 QY 721 EHNFSYPTFOAHSNNSMPGSRVQLHPPDLQSLGFGNHSIHSEDEIYENQKLAEEEE 780
 DB 721 EHNFSYPTFOAHSNNSMPGSRVQLHPPDLQSLGFGNHSIHSEDEIYENQKLAEEEE 780
 QY 781 EDNLVLTFFEDLLCFAYQVAKGMFLFKSCVHRDLAARNVLVTHGKVVKICDFGLARDIL 840
 DB 781 EDNLVLTFFEDLLCFAYQVAKGMFLFKSCVHRDLAARNVLVTHGKVVKICDFGLARDIL 840
 QY 841 SDSYVVRGNARLPVKWMAPESLFEGIYTIKSDVMSYGILLWEIFSLGVNYPYGPVDAN 900
 DB 841 SDSYVVRGNARLPVKWMAPESLFEGIYTIKSDVMSYGILLWEIFSLGVNYPYGPVDAN 900
 QY 901 FYKLIQSGFKMEQPFYATGEGIYFVMSQWAFDGRKRPSPNLTSLGCOLAABEACIRT 960
 DB 901 FYKLIQSGFKMEQPFYATGEGIYFVMSQWAFDGRKRPSPNLTSLGCOLAABEACIRT 960
 QY 961 -----SIHLPKQAAPQORGSLRAQSPQROVKIHRS 992
 DB 961 MGVNVEHPSIYQNRRLPSREAGS-EPPSPQAQVKIHRS 1000

RESULT 13
 AAR75961
 ID AAR75961 standard; Protein; 993 AA.
 XX AAR75961;
 XX 25-MAR-2003 (updated)
 DT 29-DEC-1995 (first entry)
 XX Human STK-1.
 XX STK-1; receptor PTK; protein tyrosine kinase.
 XX Homo sapiens.
 XX Key Location/Qualifiers
 FT Peptide 1..23
 FT /label= signal
 FT /note= "unique PTK receptor sequence"
 FT Cleavage-site 25..26
 FT Domain 35..524
 FT /note= "Part of extracellular domain contg.
 22 Cys residues marked on Fig. 1"
 FT Modified-site 43..45
 FT /label= potential N-linked glycosylation site
 FT Modified-site 100..102
 FT /label= see above
 FT Modified-site 151..153
 FT /label= see above
 FT Modified-site 250..252
 FT /label= see above
 FT Modified-site 306..308
 FT /label= see above
 FT Modified-site 323..325
 FT /label= see above
 FT Modified-site 351..353
 FT /label= see above
 FT Modified-site 354..356
 FT /label= see above
 FT Modified-site 473..475

FT Modified-site /label= see above
 FT 502..504
 FT /label= see above
 FT 542..562
 FT /label= transmembrane spanning region
 FT 708..782
 FT /label= kinase insert region
 FT 617..622
 FT /label= ATP binding domain
 FT 835..840
 FT /label= WMAPES motifs
 FT /note= "cytoplasmic domain"
 FT 808..813
 FT /note= "used to design PCR oligos"
 FT 870..875
 FT /note= "used to design PCR oligos"
 XX
 PN W09519175-A1.
 XX
 XX 20-JUL-1995.
 XX
 XX 06-JAN-1995; 95WO-US00176.
 XX
 XX 14-JAN-1994; 94US-0183211.
 XX (UYJO) UNIV JOHNS HOPKINS.
 PA (UYPE-) UNIV PENNSYLVANIA.
 XX
 XX Civin CI, Gewirtz AM, Small D;
 XX
 XX WPI; 1995-263709/34.
 DR N-PSDB; AAQ91356.
 XX
 XX
 PT Artificial STK-1 gene and gene-specific anti-sense oligo:nucleotide
 PT used to treat neoplastic diseases and as bone marrow purging agents for
 PT treating leukaemia and neoplasma
 XX
 PS Disclosure; Fig 1; 66pp; English.
 XX
 CC The STK-1 gene encodes a receptor PTK which is expressed in
 CC proliferating hematopoietic stem cells but not in quiescent stem
 CC cells. The STK-1 gene is also expressed in certain malignant cells
 CC of non-hematopoietic origin. An antisense oligo specific for STK-1
 CC is an oligo having a sequence (i) capable of forming a stable
 CC triplex with a portion of the STK-1 gene, or (ii) capable of forming
 CC a stable duplex with a portion of an mRNA transcript of the STK-1
 CC gene. Antisense oligos capable of forming a stable duplex with a
 CC portion of a STK-1 mRNA transcript are given in AAQ91536 FT and in
 CC AAQ91537 and AAQ91538. The antisense oligos of the invention are useful
 CC in the treatment of hematologic malignancies characterised by
 CC STK-1 expression. Several of the conserved domains of PTKs
 CC including the ATP binding domain and the WMAPES motifs are
 CC found in the STK-1 protein (see AAQ75961 FT).
 CC (Updated on 25-MAR-2003 to correct PN field.)
 XX
 SQ Sequence 993 AA;

Query Match 84.3%; Score 4436.5; DB 16; Length 993;
 Best Local Similarity 84.2%; Pred. No. 0;
 Matches 837; Conservative 57; Mismatches 89; Indels 11; Gaps 4;

QY 1 MRALAQSRDRLLLVVLSYMLETVTNQDLPVVKCVLI SHENNGSSAGKPSYSYRVRGS 60
 DB 1 MPALA-RDAGTVLLVVFSAAMIFGTITNQDLPVVKCVLIHKNNDSSVGKSSYPVMS 59
 QY 61 PEDLQTPRQSGTVYEAATVEVASGSITLQVLA TPQDLSCLVWFKHSLCCQPHD 120
 DB 60 PEDLGCALRPQSSGTVYEAATVEVASGSITLQVLA TPQDLSCLVWFKHSLCCQPHD 119
 QY 121 LQNRGVTSMAINVTQAGEYLLHQSERANTVLT VNVRTQLVLRPYPFRKMENQ 180
 DB 120 LQNRGVTSMAINVTQAGEYLLHQSERANTVLT VNVRTQLVLRPYPFRKMENQ 179

QY 181 DALLCISEGVPEPTVEVWVLCSSHRESCKEGPAVVKKEKVLHFGTDIRCCARNALGR 240
 DB 180 DALVCISEGVPEPTVEVWVLCDSQGESCKBESPAVVKKEKVLHFGTDIRCCARNELGR 239
 QY 241 ECTKLTIDLNOAPOSTLPOLFLKVGEPWIRKAIHVNHGFLTWELDKALEEGSYFE 300
 DB 240 ECTRLFTIDLNOTPTTLPOLFLKVGEPWIRKAVHVNHGFLTWELDKALEEGNYFE 299
 QY 301 MSTYSTNRTMIRILLAFVSSVGRNDGYTTCSSSKHPSQSALVTILEKGFINTSSOEY 360
 DB 300 MSTYSTNRTMIRILPAFVSSVARNDGYTTCSSSKHPSQSALVTIVEKGFINTSSSEY 359
 QY 361 EIDPYEKFCSVRFKAYPRICRTWIFSOASFCEQGLEDEYGISKECDHNKPGVIFY 420
 DB 360 EIDQYEEFCFSVRFKAYPQIRCTWTFSRKSFCEQKGLDNGYSISKPCNKHQPGVIFY 419
 QY 421 AENDDAQFTKMTNIRKPKQVLANASASOASCS SDGYPLPSWTWKCKSDKSPNCTEIEP 480
 DB 420 AENDDAQFTKMTNIRKPKQVLANASASOASCS SDGYPLPSWTWKCKSDKSPNCTEIEP 479
 QY 481 EGVNKKANRKFVGQWVSSSTLNMSEAGKGLLVKCCAYNSMGTSCTIFLNSPGPPFIQ 540
 DB 480 EGVNKRANRKFVGQWVSSSTLNMSEAIKGFVKKCCAYNSLGTSCETILLNSPGPPFIQ 539
 QY 541 DNISFYATTIGLCLPFIWLVILICHYKKOPRYESOLQMIQVTPGLDNEYFYVDFRDEY 600
 DB 540 DNISFYATTIGVCLLFIVLTLICHYKKOPRYESOLQMIQVTPGLDNEYFYVDFRDEY 599
 QY 601 DLKWEPPRENLEFGKVLGSGAFGRVNNATAYG ISKTGVSIOQAVKMLKEKADSCKEKALM 660
 DB 600 DLKWEPPRENLEFGKVLGSGAFGRVNNATAYG ISKTGVSIOQAVKMLKEKADSSERKALM 659
 QY 661 SELKMMTHLGHDNIYNLLGACTLGPVYLIP EYCCYGDLLNLYLRKREKPHRTWTIEFK 720
 DB 660 SELKMMTQLGSHENIYNLLGACTLSGP IYLIPEYCCYGDLLNLYLRKREKPHRTWTIEFK 719
 QY 721 EHNFSYPTFOAHNSSMFSGSREVQLHPDL DQLSGFNGSHSIHSEDEIEYENQKRLAEEEE 780
 DB 720 EHNFSYPTFQSHNPSMFGSREVQIHPDSD QISGLHNGSHFSEDEIEYENQKRLA---EEE 777
 QY 781 EDLNVLTFEEDLLCFAYQVAKGMEFLFKS CVHRDLAARNVLTGHGVVKICDFGLARDIL 840
 DB 778 EDLNVLTFEEDLLCFAYQVAKGMEFLFKS CVHRDLAARNVLTGHGVVKICDFGLARDIM 837
 QY 841 SDSYVVRGNARLPVKWMAPESLFEGIYTIK SDVMSYGILLWEIFSLGVNPPYGPDPVAN 900
 DB 838 SDSYVVRGNARLPVKWMAPESLFEGIYTIK SDVMSYGILLWEIFSLGVNPPYGPDPVAN 897
 QY 901 FYKLIOGFMQEPFYATGIIYFVMOQSWAFD SRKRPSPNLTSLFLGCOLAABEAC--- 957
 DB 898 FYKLIOGFMQDPFYATBEIYIIMQSWAFD SRKRPSPNLTSLFLGCOLAABEAMYQN 957
 QY 958 ----IRTSIHLPKQAAPQORG-GLRAQSPOROVK 986
 DB 958 VDRVSECPHTYQNRPRPFGRMDLGLLSPQAQVE 991

RESULT 14
 AAR67816
 ID AAR67816 standard; Protein; 993 AA.
 XX
 AC AAR67816;
 XX AC
 XX 25-MAR-2003 (updated)
 DT 18-AUG-1995 (first entry)
 XX
 XX Flk2 receptor protein-tyrosine-kinase.
 DE Human Flk2; receptor protein-tyrosine-kinase; primitive
 KW hematopoietic cell; fetal liver kinase; diagnostic ligand
 KW isolation; bone marrow disease therapy.
 XX Homo sapiens.
 OS

XX	Key	Location/Qualifiers	
PH	Peptide	1..27	
FT	Protein	/note= "signal peptide"	
FT	Domain	28...993	
FT	Domain	/note= "mature protein"	
FT	Domain	28...543	
FT	Domain	/note= "extracellular receptor domain"	
FT	Domain	544..563	
FT	Domain	/note= "transmembrane region"	
FT	Domain	564..993	
FT	Domain	/note= "intracellular catalytic domain"	
XX	WO9500554-A2.		
XX	05-JAN-1995.		
XX	17-JUN-1994;	94WO-US06944.	
XX	18-JUN-1993;	93US-0080244.	
PR	21-JUN-1993;	93US-0081508.	
PR	23-NOV-1993;	93US-0157490.	
XX	(UYPR-) UNIV PRINCETON.		
PA	Lemischka IR;		
XX	WPI; 1995-052014/07.		
DR	N-PSDB; AAQ81013.		
XX	Ligand for receptor protein tyrosine kinase - useful for the stimulation of primitive haematopoietic stem cells causing proliferation and/or differentiation		
XX	Disclosure; Fig 1b; 131pp; English.		
XX	The sequence corresponds to a human Flk2 (fetal liver kinase) receptor protein-tyrosine-kinase, which is expressed in primitive hematopoietic cells but not in mature hematopoietic cells. The protein is useful in isolation of receptor ligands, which have applications in diagnosis of bone marrow disorders and in stimulating proliferation and/or differentiation of primitive hematopoietic stem cells.		
CC	(Updated on 25-MAR-2003 to correct PN field.)		
XX	Sequence 993 AA;		
XX	Query Match 84.1%; Score 4429.5; DB 16; Length 993;		
XX	Best Local Similarity 84.1%; Pred. No. 0;		
XX	Matches 836; Conservative 57; Mismatches 90; Indels 11; Gaps 4;		
QY	1 MRALQSRDRRLLLVLSVMILETVNQDLPVKICVLIHSHENNGSSAGKPSYRMVGRS 60		
DB	1 MPALA-RDAGTVPLLVVFSAMIFGTITNQDLPVKICVLIHSHENNGSSAGKPSYRMVGRS 59		
QY	61 PEDLQTPROSGTVYEAATVEAEGSITLQVLAATPGDLSCLVFKHSSLCQHPD 120		
DB	60 PEDLGCALRPQSSGTVYEAATVEAEGSITLQVLAATPGDLSCLVFKHSSLCQHPD 119		
QY	121 LQNRGIVSMALNVTVTQAGEYLLIHOSERANYTLVTVNVRDQTLVYLRPRPKMENQ 180		
DB	120 LQNRGVSMVILKMTQAGEYLLFIOSEATNYTLFTVIRNTLLTYLRPRPKMENQ 179		
QY	181 DALLCISEGVPETVENVLCSHRESCKEKGPAVVRKEEVLHFGTDIRCCARNALGR 240		
DB	180 DALVCISESVPETVENVLCSHRESCKEKGPAVVRKEEVLHFGTDIRCCARNALGR 239		
QY	241 ECTKLTIDLNQAPSTLPFLKVGEPPLWIRCKAHVNHGFLTWLEDKALEEGSYFE 300		
DB	240 ECTRLFTIDLNQAPSTLPFLKVGEPPLWIRCKAHVNHGFLTWLEDKALEEGSYFE 299		
QY	301 MSTYSTNRTMIRILLAFVSSVGRNDTGYTCCSSKHPQSALVTLTKGFINATSSQEEY 360		

DB	300	MSTYSTNRTMIRILLAFVSSVGRNDTGYTCCSSKHPQSALVTLTKGFINATSSQEEY 359
QY	361	EIDPYEKECFSVRFKAYPRIRCTNIPSOASFPCEQRLGDEGYSISKFCDHQKHPGYIFY 420
DB	360	EIDQYEEFCFSVRFKAYPQIRCTNIPSOASFPCEQRLGDEGYSISKFCNHKHPGYIFY 419
QY	421	AENDDAQFTKFTLNIRKKPOVLNANASQASCSDDGYPVLPSTWTKKCKSDKSPNCTEIP 480
DB	420	AENDDAQFTKFTLNIRKKPOVLNANASQASCSDDGYPVLPSTWTKKCKSDKSPNCTEIP 479
QY	481	EGVWNKKANRVFGQWSSSTLNMSSEAGKLLVKKCAYNMGTSCETIFINSPGPPFIQ 540
DB	480	EGVWNKKANRVFGQWSSSTLNMSSEAGKLLVKKCAYNMGTSCETIFINSPGPPFIQ 539
QY	541	DNISFYATIGLCLPFIIVLVIIVLICHYKQKQRYESQLOMIQVTPGLDNEFYVDFRDY 600
DB	540	DNISFYATIGLCLPFIIVLVIIVLICHYKQKQRYESQLOMIQVTPGLDNEFYVDFRDY 599
QY	601	DLKWEFPRENLEFGKVLGSGAFGRVMNATAYGISKTGVSIOAVKMLKEKADSCKEALM 660
DB	600	DLKWEFPRENLEFGKVLGSGAFGRVMNATAYGISKTGVSIOAVKMLKEKADSCKEALM 659
QY	661	SELKQMTLGHHDNIIVNLGACTLGGPIYLIFEYCCYGDLLNLYRSKREKPHRTWTEIFK 720
DB	660	SELKQMTLGHHDNIIVNLGACTLGGPIYLIFEYCCYGDLLNLYRSKREKPHRTWTEIFK 719
QY	721	EHNFSYPTFOAHNSNMPGSRREVQLHPDLQSGFNNGSIHSEDEIYEYENOKRLAEE 780
DB	720	EHNFSYPTFOAHNSNMPGSRREVQLHPDLQSGFNNGSIHSEDEIYEYENOKRLAEE 777
QY	781	EDNLVATFEDLLCFAYQVAKGMEFLPKSCVHRDLAARNVLTGKGVKICDFGLARDIL 840
DB	778	EDNLVATFEDLLCFAYQVAKGMEFLPKSCVHRDLAARNVLTGKGVKICDFGLARDIL 837
QY	841	SDSYVVRGNARLPVKWMAPESLFEGYITIKSDVMSYGILLWEIFSLGVPYGPVVDAN 900
DB	838	SDSYVVRGNARLPVKWMAPESLFEGYITIKSDVMSYGILLWEIFSLGVPYGPVVDAN 897
QY	901	FYKLIQSGFKMEQPFYATEGIYFVMSQWAFDSRKRPSFNLTSFLGQCLAEAEAC--- 957
DB	898	FYKLIQSGFKMEQPFYATEGIYFVMSQWAFDSRKRPSFNLTSFLGQCLAEAEAMQN 957
QY	958	----IRTSIHLPKQAAPQORG-GLRAQSPQKQV 986
DB	958	VDGRVSECPHTYQNRPRPFRSREMDLGLLSPQAQVE 991
XX	RESULT 15	
XX	AAR67536	
XX	ID AAR67536 standard; Protein; 993 AA.	
XX	AAR67536;	
XX	AC AC	
XX	XX XX	
DT	25-MAR-2003 (updated)	
DT	04-JUN-1995 (first entry)	
XX	XX	
DE	Human flk-2.	
XX	XX	
KW	Fetal liver kinase-2; flk-2; protein tyrosine-kinase receptor;	
KW	hematopoiesis; stem cell.	
XX	XX	
OS	Homo sapiens.	
XX	XX	
PH	Key	Location/Qualifiers
FT	Peptide	1..27
FT	Domain	/label= sig_peptide
FT	Domain	28..543
FT	Region	/label= Extracellular_receptor_domain
FT	Region	544..563
FT	Domain	/label= Transmembrane_region
FT	Domain	564..993
FT	Domain	/label= Intracellular_catalytic_domain
XX	XX	

PN US5367057-A.
 XX 22-NOV-1994.
 XX 30-APR-1993; 93US-0055269.
 XX 02-APR-1991; 91US-0679666.
 XX 28-JUN-1991; 91US-0728913.
 XX 15-NOV-1991; 91US-0793065.
 XX 26-DEC-1991; 91US-0813593.
 XX 26-JUN-1992; 92US-0906397.
 XX 12-NOV-1992; 92US-0975049.
 XX 19-NOV-1992; 92US-0977451.
 XX 30-APR-1993; 93US-0055269.
 XX (UYPR-) UNIV PRINCETON.
 XX Lemischka IR;
 XX WPI; 1995-005894/01.
 XX N-PSDB; AAQ79069.
 XX Murine flk-2 receptor protein tyrosine kinase - used to stimulate
 PT proliferation and/or stimulation of primitive mammalian
 PT haematopoietic stem cells in vitro or in vivo.
 XX Disclosure; Fig. 2A-1F; 69pp; English.
 XX cDNAs encoding receptor protein tyrosine-kinases, mouse fetal liver
 CC kinase-2 (flk-2), human flk-2 and mouse flk-1 are given in AAQ79068-70,
 CC respectively, and the deduced amino acid sequences in AAQ67535-37,
 CC respectively.
 CC (Updated on 25-MAR-2003 to correct PF field.)
 XX Sequence 993 AA;
 SQ

Query Match 84.1%; Score 4429.5; DB 16; Length 993;
 Best Local Similarity 84.1%; Pred. No. 0;
 Matches 836; Conservative 57; Mismatches 90; Indels 11; Gaps 4;

Qy 1 MRALAQRSDRLLLVLSVMILETVTNQDLPVKVLISHENNGSSAGKPSYRMVRGS 60
 Db 1 MPALA-RDAGTVPLLVFSANI FGTITNQDLPVKVLINKNNDSSVGKSSYPWVSES 59
 Qy 61 PEDLOCTPRQSEGTVEAATVEVAESGSIITLQVOLATPGDLSCLWFKHSSLCQPHFD 120
 Db 60 PEDLGCALRPOS SGTVEA AAVEVDVSASITLQVLVDAPGNISCLWVFKHSSLNCQPHFD 119
 Qy 121 LQNRGIYSMAILNTTQAGEYILHIQSERANYTVLFTVVRDTQLYVLRPPYFRKMENQ 180
 Db 120 LQNRGVVSMVLKMTQAGEYLLFIQSEATNYTILFTVSRINTLLYTLRRPYFRKMENQ 179
 Qy 181 DALLCISEGVPEPTVEWVLCSSHRESCKECPAVVRKEEVLHELFGTDIRCCARNALGR 240
 Db 180 DALVCISEVPEPIVEWVLCDSQESCKEESPAVVRKEEVLHELFGTDIRCCARNELGR 239
 Qy 241 ECTKLFITDLNQAQSTLPQLFLKVGELPWIRCKAIHVNHGFGTLWELEDKALBEGSYFE 300
 Db 240 ECTRLFTIDLNTQTPTTLPLQLFLKVGELPWIRCKAVHNVHFGTLWELENKALBEGSYFE 299
 Qy 301 MSTYSTNRTMIRILLAFVSSVGRNDTGYYTSSSKHPQSQSALVTILEKGFNATSSOEY 360
 Db 300 MSTYSTNRTMIRILLAFVSSVARNDTGYYTSSSKHPQSQSALVTIVGKGFNATNSSEY 359
 Qy 361 EIDPYEFCFSVRKAYPRICRTWIFSOASPCBORGLEDGYSISKPCDHKNKPGEXIFY 420
 Db 360 EIDQYEEFCFSVRKAYPQIRCTWTFRKSPFCQKGLDNGYSISKPCNHKHQGEYIFH 419
 Qy 421 AENDDAQFTKMTNIRKKQVLANASQASCSGSGYPLPSWTWKKCSKSPNCTBEIP 480
 Db 420 AENDDAQFTKMTNIRKKQVLAESASQASCSGSGYPLPSWTWKKCSKSPNCTBEIT 479
 Qy 481 EGVWNRKANRKVFCQWVSSSTLANSEAGKGLLVKCCAYNSMGTSCETIFLNSPGPPFIQ 540

Db 480 EGVWNRKANRKVFCQWVSSSTLANSEAIKGLVLKCCAYNSMGTSCETIFLNSPGPPFIQ 539
 Qy 541 DNISFYATIGLCLPFIIVLVI CHYKQPRYBSQLQMIQVTPGLDNEYFYVDFRDY 600
 Db 540 DNISFYATIGVCLLFIIVLTLI CHYKQPRYBSQLQMIQVTPGLDNEYFYVDFRDY 599
 Qy 601 DLKWEFFPRENLEPQVLGSGAFGRVMNATAYIGISKTGVSIOAVVMKLEKADSCKEALM 660
 Db 600 DLKWEFFPRENLEPQVLGSGAFGRVMNATAYIGISKTGVSIOAVVMKLEKADSCKEALM 659
 Qy 661 SELKQMTLGHHDNI VNLGACTLSGPVYLIFEYCCYGDLLNLYLSRKRKHRTWTEIFK 720
 Db 660 SELKQMTLGHHDNI VNLGACTLSGPVYLIFEYCCYGDLLNLYLSRKRKHRTWTEIFK 719
 Qy 721 EHNFSYPTFOAHNSNSMPSRREVQLHPPLDQLSGFNGNSIHSDEIEYENQKRLAEBEE 780
 Db 720 EHNFSYPTFOAHNSNSMPSRREVQLHPPLDQLSGFNGNSIHSDEIEYENQKRLAEBEE 777
 Qy 781 EDLNVLTPEDLCLCPAYQVAKGMEFLEPKSCVHRDLAARNVLVTRGVKVKICDFGLARDIL 840
 Db 778 EDLNVLTPEDLCLCPAYQVAKGMEFLEPKSCVHRDLAARNVLVTRGVKVKICDFGLARDIM 837
 Qy 841 SDSSVVRGNARLPVKWMAPELSEFEGYITIKSDVMSYGILWEIFSLGVNYPGIPVDAN 900
 Db 838 SDSSVVRGNARLPVKWMAPELSEFEGYITIKSDVMSYGILWEIFSLGVNYPGIPVDAN 897
 Qy 901 FYKLIQSGFKMEQPFYATEGIYFVMQSCWAFDSRKRPSFPNLTSPFLGCOLAEAEAC--- 957
 Db 898 FYKLIQSGFKMEQPFYATEGIYFVMQSCWAFDSRKRPSFPNLTSPFLGCOLAEAEAMYN 957
 Qy 958 ----IRTSIHLPKQAAPQORG-GLRAQSPQRQVK 986
 Db 958 VDGVRSECPHTYQNRPFPSREMDLGLLSPOAQVE 991

Search completed: August 26, 2003, 07:24:09
 Job time : 67.4685 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2003, 07:27:59 ; Search time 45.4771 Seconds
(without alignments)
2878.506 Million cell updates/sec

Title: US-09-919-408A-2
Perfect score: 5264
Sequence: 1 MRALAQSRDLRLVLSV.....RGGLRAQSPQVKIHRS 992

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 497079 seqs, 131961718 residues

Total number of hits satisfying chosen parameters: 497079

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*

1:	/cgn2_5/ptodata/1/pubaa/US07_PUBCOMB.pep.*
2:	/cgn2_6/ptodata/1/pubaa/PCT_NEW_PUB.pep.*
3:	/cgn2_6/ptodata/1/pubaa/US06_PUBCOMB.pep.*
4:	/cgn2_6/ptodata/1/pubaa/US06_PUBCOMB.pep.*
5:	/cgn2_6/ptodata/1/pubaa/US07_NEW_PUB.pep.*
6:	/cgn2_6/ptodata/1/pubaa/PCTUS_PUBCOMB.pep.*
7:	/cgn2_6/ptodata/1/pubaa/US08_NEW_PUB.pep.*
8:	/cgn2_6/ptodata/1/pubaa/US08_PUBCOMB.pep.*
9:	/cgn2_6/ptodata/1/pubaa/US09_PUBCOMB.pep.*
10:	/cgn2_6/ptodata/1/pubaa/US09B_PUBCOMB.pep.*
11:	/cgn2_6/ptodata/1/pubaa/US09C_PUBCOMB.pep.*
12:	/cgn2_6/ptodata/1/pubaa/US09_NEW_PUB.pep.*
13:	/cgn2_6/ptodata/1/pubaa/US10_PUBCOMB.pep.*
14:	/cgn2_6/ptodata/1/pubaa/US10B_PUBCOMB.pep.*
15:	/cgn2_6/ptodata/1/pubaa/US10C_PUBCOMB.pep.*
16:	/cgn2_6/ptodata/1/pubaa/US10_NEW_PUB.pep.*
17:	/cgn2_6/ptodata/1/pubaa/US60_NEW_PUB.pep.*
18:	/cgn2_6/ptodata/1/pubaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5264	100.0	992	9	US-09-919-408-2
2	5264	100.0	992	10	US-09-872-136-2
3	4429.5	84.1	993	9	US-09-919-408-4
4	4429.5	84.1	993	10	US-09-872-136-4
5	1223	23.2	975	15	US-10-192-867-2
6	1221.5	23.2	972	10	US-09-944-807-10
7	1196	22.7	976	14	US-10-099-895-32
8	1196	22.7	976	15	US-10-192-867-4
9	1152.5	21.9	1088	11	US-09-961-403-4
10	1152.5	21.9	1089	9	US-09-769-987-2
11	1152.5	21.9	1089	10	US-09-919-497-90
12	1152.5	21.9	1089	10	US-09-866-510-2
13	1150.5	21.9	1089	10	US-09-955-363-36
14	1149.5	21.8	1089	10	US-09-866-510-10
15	1148.5	21.8	1089	10	US-09-866-510-4

16	1147.5	21.8	1089	10	US-09-866-510-8
17	1146.5	21.8	1089	10	US-09-866-510-6
18	1078	20.5	1090	10	US-09-866-510-14
19	1078	20.5	1106	10	US-09-866-510-22
20	1078	20.5	1106	10	US-09-955-363-2
21	1074	20.4	1106	10	US-09-866-510-16
22	1073	20.4	1106	10	US-09-866-510-20
23	1072	20.4	1106	10	US-09-866-510-18
24	998.5	19.0	1338	12	US-10-021-660-84
25	998.5	19.0	1338	15	US-10-059-585-44
26	961.5	18.3	1356	15	US-10-022-939-2
27	961.5	18.3	1356	15	US-10-100-405A-2
28	960.5	18.2	1356	11	US-09-969-037-7
29	954.5	18.1	1298	10	US-09-982-610-33
30	954.5	18.1	1298	15	US-10-262-538-32
31	954.5	18.1	1354	15	US-10-262-538-30
32	954.5	18.1	1363	11	US-09-375-248-2
33	952.5	18.1	1368	15	US-10-105-901-34
34	952	18.1	1363	11	US-09-375-248-19
35	943.5	17.9	1362	15	US-10-105-901-33
36	941	17.9	386	9	US-09-939-754-6
37	941	17.9	386	10	US-09-939-832-6
38	941	17.9	386	10	US-09-939-833-6
39	937.5	17.8	1367	9	US-09-766-678-2
40	936.5	17.8	1363	15	US-10-105-901-32
41	926.5	17.6	1367	9	US-09-919-408-6
42	926.5	17.6	1367	10	US-09-872-136-6
43	862	16.4	367	9	US-09-939-754-9
44	862	16.4	367	10	US-09-939-832-9
45	862	16.4	367	10	US-09-939-833-9

ALIGNMENTS

RESULT 1

US-09-919-408-2
; Sequence 2, Application US/09919408
; Patent No. US20020072077A1
; GENERAL INFORMATION:
; APPLICANT: Lemischka, Ihor R.
; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL
; RECEPTORS AND THEIR LIGANDS
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ImClone Systems Incorporated
; STREET: 180 Varick Street
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10014
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/919,408
; FILING DATE: 31-Jul-2001
; CLASSIFICATION: <Unknown>
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: 07/977,451
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US 07/906,397
; FILING DATE: 26-JUN-1992
; APPLICATION NUMBER: US PCT/US92/05401
; FILING DATE: 26-JUN-1992
; APPLICATION NUMBER: TW 81102961
; FILING DATE: 15-APR-1992
; APPLICATION NUMBER: US PCT/US92/02750
; FILING DATE: 02-APR-1992
; APPLICATION NUMBER: US 07/813,593
; FILING DATE: 24-DEC-1991


```

; APPLICATION NUMBER: US 07/793,065
; FILING DATE: 15-NOV-1991
; APPLICATION NUMBER: US 07/728,913
; FILING DATE: 28-JUN-1991
; APPLICATION NUMBER: US 07/679,666
; FILING DATE: 02-APR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Feit, Irving N.
; REGISTRATION NUMBER: 28,601
; REFERENCE/DOCKET NUMBER: LEM-3-7P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-645-1405
; TELEFAX: 212-645-2054
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 992 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-919-408-2

```

```

Query Match 100.0%; Score 5264; DB 9; Length 992;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 992; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MRALAQRSDRLLLVLSVMILETVNODLPVTKVLI SHENNGSSAGKPSRYMVRGS 60
Db 1 MRALAQRSDRLLLVLSVMILETVNODLPVTKVLI SHENNGSSAGKPSRYMVRGS 60

Qy 61 PEDLQCTPRQSEGTVEAATVEAESGSIITQVQLATPGDLSCLWFKHSSLCQPHFD 120
Db 61 PEDLQCTPRQSEGTVEAATVEAESGSIITQVQLATPGDLSCLWFKHSSLCQPHFD 120

Qy 121 LQNRGIVSMALNTETQAGEYLLHIQSERANYTVLTVNVRTQYVLRPPYPRKMNQ 180
Db 121 LQNRGIVSMALNTETQAGEYLLHIQSERANYTVLTVNVRTQYVLRPPYPRKMNQ 180

Qy 181 DALLCISEGPEPTVEVWLCSSHRESCKECPAVVRKEEKVHLFGTDIRCCARNALGR 240
Db 181 DALLCISEGPEPTVEVWLCSSHRESCKECPAVVRKEEKVHLFGTDIRCCARNALGR 240

Qy 241 ECTKLFITDLNQAPQSTLPQLFLKVGEPPLIRCKAIHVNHGFGITWLEDKALBEGSYFE 300
Db 241 ECTKLFITDLNQAPQSTLPQLFLKVGEPPLIRCKAIHVNHGFGITWLEDKALBEGSYFE 300

Qy 301 MSTYSTNRTMIRILLAFVSSVGRNDTGYTTCSSSKHPSQSALVTILEKGFNATSSQEEY 360
Db 301 MSTYSTNRTMIRILLAFVSSVGRNDTGYTTCSSSKHPSQSALVTILEKGFNATSSQEEY 360

Qy 361 EIDPYEKFCFSVRPKAYPRIRCTWIPSOAGPCCPQORGLDGYSTSKPCDHKNKPEYIFY 420
Db 361 EIDPYEKFCFSVRPKAYPRIRCTWIPSOAGPCCPQORGLDGYSTSKPCDHKNKPEYIFY 420

Qy 421 AENDDAQFTKMTLINIRKKQVLANASASQSSDGYPLPSPWTKKCSKSPNCTEIP 480
Db 421 AENDDAQFTKMTLINIRKKQVLANASASQSSDGYPLPSPWTKKCSKSPNCTEIP 480

Qy 481 EGWNKANKRVQGWVSSSTLNMSEAGKLLVKCCAYNSMGTSCTETFLNSPGPPFFIQ 540
Db 481 EGWNKANKRVQGWVSSSTLNMSEAGKLLVKCCAYNSMGTSCTETFLNSPGPPFFIQ 540

Qy 541 DNTSFYATIGLCUPFIVLVILCHYKQFRYESQLQMTQVTPGLDNEYFYVDFRDEY 600
Db 541 DNTSFYATIGLCUPFIVLVILCHYKQFRYESQLQMTQVTPGLDNEYFYVDFRDEY 600

Qy 601 DLKWEFPRENLEFGKVLGSAFGVNMATAYGTSKTVSQAVALMKLKEKADSCKEALM 660
Db 601 DLKWEFPRENLEFGKVLGSAFGVNMATAYGTSKTVSQAVALMKLKEKADSCKEALM 660

Qy 661 SELKMMTHLGHNDINVLNLAGCTLSGPVYLIFCYCCYGDLLNLYRSKREKPHRTWTBIFK 720
Db 661 SELKMMTHLGHNDINVLNLAGCTLSGPVYLIFCYCCYGDLLNLYRSKREKPHRTWTBIFK 720

```

```

RESULT 2
US-09-872-136-2
; Sequence 2, Application US/09872136
; Patent No. US20020119545A1
; GENERAL INFORMATION:
; APPLICANT: Lemischka, Ihor R.
; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL
; RECEPTORS AND THEIR LIGANDS
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: InClone Systems Incorporated
; STREET: 180 Varick Street
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10014
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/872,136
; FILING DATE: 01-Jun-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/208,786
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US/09/021,324
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US/07/977,451
; FILING DATE: 1992-11-19
; APPLICATION NUMBER: US 07/906,397
; FILING DATE: 26-JUN-1992
; APPLICATION NUMBER: US PCT/US92/05401
; FILING DATE: 26-JUN-1992
; APPLICATION NUMBER: TW 81102961
; FILING DATE: 15-APR-1992
; APPLICATION NUMBER: US PCT/US92/02750
; FILING DATE: 02-APR-1992
; APPLICATION NUMBER: US 07/813,593
; FILING DATE: 24-DEC-1991
; APPLICATION NUMBER: US 07/793,065
; FILING DATE: 15-NOV-1991
; APPLICATION NUMBER: US 07/728,913
; APPLICATION NUMBER: US 07/679,666
; FILING DATE: 02-APR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Feit, Irving N.
; REGISTRATION NUMBER: 28,601

```

```

Qy 721 EHNFSYPTFOAHNSNMPGSRVQLHPPDLQSLGFGNGNSHSEDEIEYENOKRLAESEE 780
Db 721 EHNFSYPTFOAHNSNMPGSRVQLHPPDLQSLGFGNGNSHSEDEIEYENOKRLAESEE 780
Qy 781 EDNLVLTFFEDLLCFAYQVAKGMEFLPKSCVHRDLAARNVLVTHGKVKICDFGLARDIL 840
Db 781 EDNLVLTFFEDLLCFAYQVAKGMEFLPKSCVHRDLAARNVLVTHGKVKICDFGLARDIL 840
Qy 841 SDSSYVVRGNARLPVKWMAPESLPEGIYTIKSDVMSYGILLWEIFSLGVNPPGIPVDAN 900
Db 841 SDSSYVVRGNARLPVKWMAPESLPEGIYTIKSDVMSYGILLWEIFSLGVNPPGIPVDAN 900
Qy 901 FYKLIQSGFKMEQPFYATEGIYFVMSQWAFDSKRPSFNNLTGFLGCOLAEAEACIRT 960
Db 901 FYKLIQSGFKMEQPFYATEGIYFVMSQWAFDSKRPSFNNLTGFLGCOLAEAEACIRT 960
Qy 961 SIHLPKQAAPQORGLRAQSPQROVKIHRRS 992
Db 961 SIHLPKQAAPQORGLRAQSPQROVKIHRRS 992

```

REFERENCE/DOCKET NUMBER: LEM-3-7P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-645-1405
TELEFAX: 212-645-2054
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 992 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-872-136-2

Query Match 100.0%; Score 5264; DB 10; Length 992;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 992; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRALQSRDRRLLLVLSVMIETVTDNODLVKVLISHENNGSSAGKPSYSYMRGS 60
DB 1 MRALQSRDRRLLLVLSVMIETVTDNODLVKVLISHENNGSSAGKPSYSYMRGS 60
QY 61 PEDLOCTPRQSEGVYEATVEAEGSSITLQVOLATPGDLSCLWVFKHSLGQPHFD 120
DB 61 PEDLOCTPRQSEGVYEATVEAEGSSITLQVOLATPGDLSCLWVFKHSLGQPHFD 120
QY 121 LQNRGIVSMAILNVETQAGEYLLHIQSERANYTVLFTVNVVDTQLYVLRPFYFRMENQ 180
DB 121 LQNRGIVSMAILNVETQAGEYLLHIQSERANYTVLFTVNVVDTQLYVLRPFYFRMENQ 180
QY 181 DALLCISEGVPEPTVEWVLCSSHRSCKEGBPAVRKEKVLHELPGTDIRCARNALGR 240
DB 181 DALLCISEGVPEPTVEWVLCSSHRSCKEGBPAVRKEKVLHELPGTDIRCARNALGR 240
QY 241 ECKTLFTIDLNAQPOSTLPOLFKVGEPLWIRCKAIHVNHGGLTWELDKALEGSYFE 300
DB 241 ECKTLFTIDLNAQPOSTLPOLFKVGEPLWIRCKAIHVNHGGLTWELDKALEGSYFE 300
QY 301 MSTYSTNRTMIRILLAFVSSVGRNDGYTCTSSSKHPSQSALVTILEKGFINATSQEY 360
DB 301 MSTYSTNRTMIRILLAFVSSVGRNDGYTCTSSSKHPSQSALVTILEKGFINATSQEY 360
QY 361 EIDPYEKFCSVRFKAYPRICRTWIFSQASFPCEQRLGEGYSISKFDHKHKNPGYIFY 420
DB 361 EIDPYEKFCSVRFKAYPRICRTWIFSQASFPCEQRLGEGYSISKFDHKHKNPGYIFY 420
QY 421 AENDDAQFTMTLNRKPKQVLANASQASQSSDGYPLPSWTWKCKSDKSPNCTEIP 480
DB 421 AENDDAQFTMTLNRKPKQVLANASQASQSSDGYPLPSWTWKCKSDKSPNCTEIP 480
QY 481 EGVNKKANRKFVGQVSSSTLNMSEAGKGLVKKCCAYNSMGTSCTIFLNSPGPPFIQ 540
DB 481 EGVNKKANRKFVGQVSSSTLNMSEAGKGLVKKCCAYNSMGTSCTIFLNSPGPPFIQ 540
QY 541 DNISFYATIGLCPFTVLVLVILCHYKQFYESQLOMIQVTPLDNNEYFYVDFRDY 600
DB 541 DNISFYATIGLCPFTVLVLVILCHYKQFYESQLOMIQVTPLDNNEYFYVDFRDY 600
QY 601 DLKWEPPRENLEFGKVLGSGAFGRVNNATAYGTSKTVSIQVAVMLKEKADSCKEALM 660
DB 601 DLKWEPPRENLEFGKVLGSGAFGRVNNATAYGTSKTVSIQVAVMLKEKADSCKEALM 660
QY 661 SELKMMTHLGHNDNIWNLGACTLGPVVLIFPYCCVGDLLNVLRSKREKFRHTWTEIFK 720
DB 661 SELKMMTHLGHNDNIWNLGACTLGPVVLIFPYCCVGDLLNVLRSKREKFRHTWTEIFK 720
QY 721 EHNFSYPTFOAHNSMFGSREVQLHPPDLQSLGFGNGSIHSEDEIYENQKRLAE 780
DB 721 EHNFSYPTFOAHNSMFGSREVQLHPPDLQSLGFGNGSIHSEDEIYENQKRLAE 780
QY 781 EDNLNLTFTEDLLCFAYQVAKGMFLEFKSCVHRDLAARNVLVTHGKVKIKCDFGLARDIL 840
DB 781 EDNLNLTFTEDLLCFAYQVAKGMFLEFKSCVHRDLAARNVLVTHGKVKIKCDFGLARDIL 840

QY 841 SDSSYVVRGNARLPVKWNPESLPEGIYTIKSDVMSYGILLWEIFSLGWNPIPGIPVDAN 900
DB 841 SDSSYVVRGNARLPVKWNPESLPEGIYTIKSDVMSYGILLWEIFSLGWNPIPGIPVDAN 900
QY 901 FYKLIQSGFKMEQPFYATEGYFVQMSQWAFDSRKRPSFPNLTSLGCOLAEAEACIRT 960
DB 901 FYKLIQSGFKMEQPFYATEGYFVQMSQWAFDSRKRPSFPNLTSLGCOLAEAEACIRT 960
QY 961 SIHLPKQAAPQORGGGLRAQSPORQVKIHRERS 992
DB 961 SIHLPKQAAPQORGGGLRAQSPORQVKIHRERS 992

RESULT 3
US-09-919-408-4
Sequence 4, Application US/09919408
Patent No. US20020072077A1
GENERAL INFORMATION:
APPLICANT: Lemischka, Ihor R.
TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL RECEPTORS AND THEIR LIGANDS
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESS: ImClone Systems Incorporated
STREET: 180 Varick Street
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10014
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/919,408
FILING DATE: 31-Jul-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/977,451
FILING DATE: <Unknown>
APPLICATION NUMBER: US 07/906,397
FILING DATE: 26-JUN-1992
APPLICATION NUMBER: US PCT/US92/05401
FILING DATE: 26-JUN-1992
APPLICATION NUMBER: TW 81102961
FILING DATE: 15-APR-1992
APPLICATION NUMBER: US PCT/US92/02750
FILING DATE: 02-APR-1992
APPLICATION NUMBER: US 07/813,593
FILING DATE: 24-DEC-1991
APPLICATION NUMBER: US 07/793,065
FILING DATE: 15-NOV-1991
APPLICATION NUMBER: US 07/728,913
FILING DATE: 28-JUN-1991
APPLICATION NUMBER: US 07/679,666
FILING DATE: 02-APR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Felt, Irving N.
REGISTRATION NUMBER: 28,601
REFERENCE/DOCKET NUMBER: LEM-3-7P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-645-1405
TELEFAX: 212-645-2054
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 993 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-919-408-4

Query Match 84.1%; Score 4429.5; DB 9; Length 993;
 Best Local Similarity 84.1%; Pred. No. 0;
 Matches 836; Conservative 57; Mismatches 90; Indels 11; Gaps 4;

QY 1 MRALAORSRRLLLVLSVMIETVNDLPVKVLI SHENNGSSAGKPSVYRVGRS 60
 DB 1 MPALA-RDAGTVPLLVVFSAMIFGTINQDLPVKVLIHNKNDSSVGKSSYPMVSES 59

QY 61 PEDLQCTPRQSGTVVEAATVEAESGSIITLQVQLATPGDLSCLVWFKHSSLCQPHFD 120
 DB 60 PEDLGCALRQSSGTVVEAARVEDVDSASITLQVLDAPGNISCLVWFKHSSLCQPHFD 119

QY 121 LQNRGIVSMALNVETQAGEYLHIOSEARANYTLFTVNVRTQTLVLRPRYRKMENQ 180
 DB 120 LQNRGVVSMVLKMTQAGEYLLFIOSEATNYTLFTVIRNTLLYTLRRPYRKMENQ 179

QY 181 DALLCISEGVPEPTVEVLCSHRESCKEESGPAVVRKEEVHLFGTDIRCCARNALGR 240
 DB 180 DALVCISESVPEPIVEMVLCDQESCKEESPAVVRKEEVHLFGTDIRCCARNALGR 239

QY 241 ECTKLFITDLNQAQSTLPOLFLKVGEPILWIRCKAIIHNVHGFGLTWELDKALEEGSYFE 300
 DB 240 ECTRLFTIDLNQPTQTLPOLFLKVGEPILWIRCKAVHNVHGFGLTWELDKALEEGSYFE 299

QY 301 MSTYSTNRTMIRILLAFVSSVGRNDTGYTTCSSKHPQSALVTILEKGFINATSSQEEY 360
 DB 300 MSTYSTNRTMIRILFAPVSSVARNDTGYTTCSSKHPQSALVTIVKGFINATSSSEY 359

QY 361 EIDPYEKFCSVRFKAYPRCTRWIFQASFPCEORGLDGYISKECDHKNRGEYIF 420
 DB 360 EIDQYEFECFSVRFKAYPOIRCTWTFKSPFPCEQKGLDNGYSKFCNKHQGEYIFH 419

QY 421 AENDDAQFTKMTFLNIRKQOVLANASOASGSSDGYPLPSPWTKKCDKSPNCTEIP 480
 DB 420 AENDDAQFTKMTFLNIRKQOVLANASOASGSDGYPLPSPWTKKCDKSPNCTEIT 479

QY 481 EGVNWKANKRVFGQWSSSTLANMSEAGLLVKCCAYNSMGTSCTIFLNSPGPPFFIQ 540
 DB 480 EGVNWKANKRVFGQWSSSTLANMSEAIKGLVKCCAYNSLGTSCETILLNSPGPPFFIQ 539

QY 541 DNTSFATIGLCIPFIVLVILCHYKQFRESQLOMLOVQTGPLNEFYVDPRY 600
 DB 540 DNTSFATIGVCLLFIIVLLTLLCHYKQFRESQLOMLOVQTGSSDNEFYVDPRY 599

QY 601 DLKWEPRENLEFGKVLGSGAFGRVMNATAYISKTGVSQVAVKMLKEKADSCKEALM 660
 DB 600 DLKWEPRENLEFGKVLGSGAFGRVMNATAYISKTGVSQVAVKMLKEKADSSERALM 659

QY 661 SELKMTHTLGHNDINVLGACTLSGPVYLIFEYCCYGDLLNLYLRKREKPHRTWTEIFK 720
 DB 660 SELKMTHTLGHNDINVLGACTLSGPVYLIFEYCCYGDLLNLYLRKREKPHRTWTEIFK 719

QY 721 EHNFSYPTTQAHNSNMPGSRVQLHPPDLQSLGFGNHSIHSEDEIYEYENQKLABE 780
 DB 720 EHNFSYPTTQSHNSNMPGSRVQLHPPDLQSLGFGNHSIHSEDEIYEYENQKLABE 777

QY 781 EDNLVLTFFDLCLFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVKVCDLGLARDIL 840
 DB 778 EDNLVLTFFDLCLFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVKVCDLGLARDIM 837

QY 841 SDSYVVRGNARLPVKWMAPESLPEGIYTIKSDVMSYGILLWEIFSLGVNYPGIPVDAN 900
 DB 838 SDSYVVRGNARLPVKWMAPESLPEGIYTIKSDVMSYGILLWEIFSLGVNYPGIPVDAN 897

QY 901 FYKLIQSGFQWQPPYATEGIYFMQSCWAFDSKRPSFNLTSLFLCQLABAEAC--- 957
 DB 898 FYKLIQSGFQWQPPYATEGIYFMQSCWAFDSKRPSFNLTSLFLCQLABAEAC--- 957

QY 958 ----IRTSIHLKQAAAPQQRG-GLRAQSPORQVK 986
 DB 958 VDRGVSECPHTYQNRPRFPRMDLGLLSPOAQQVE 991

RESULT 4
 US-09-872-136-4
 ; Sequence 4, Application US/09872136
 ; Patent No. US20020119545A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Lemischka, Ihor R.
 ; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL
 ; RECEPTORS AND THEIR LIGANDS
 ; NUMBER OF SEQUENCES: 10
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: ImClone Systems Incorporated
 ; CITY: New York
 ; STATE: New York
 ; COUNTRY: U.S.A.
 ; ZIP: 10014
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/872,136
 ; FILING DATE: 01-Jun-2001
 ; CLASSIFICATION: <Unknown>
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/208,786
 ; FILING DATE: <Unknown>
 ; APPLICATION NUMBER: US/09/021,324
 ; FILING DATE: <Unknown>
 ; APPLICATION NUMBER: US/07/977,451
 ; FILING DATE: 1992-11-19
 ; APPLICATION NUMBER: US 07/906,397
 ; FILING DATE: 26-JUN-1992
 ; APPLICATION NUMBER: US PCT/US92/05401
 ; FILING DATE: 26-JUN-1992
 ; APPLICATION NUMBER: TW 81102961
 ; FILING DATE: 15-APR-1992
 ; APPLICATION NUMBER: US PCT/US92/02750
 ; FILING DATE: 02-APR-1992
 ; APPLICATION NUMBER: US 07/813,593
 ; FILING DATE: 24-DEC-1991
 ; APPLICATION NUMBER: US 07/793,065
 ; FILING DATE: 15-NOV-1991
 ; APPLICATION NUMBER: US 07/728,913
 ; FILING DATE: 28-JUN-1991
 ; APPLICATION NUMBER: US 07/679,666
 ; FILING DATE: 02-APR-1991
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Feit, Irving N.
 ; REGISTRATION NUMBER: 28,601
 ; REFERENCE/DOCKET NUMBER: LEM-3-7P
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 212-645-1405
 ; TELEFAX: 212-645-2054
 ; INFORMATION FOR SEQ ID NO: 4:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 993 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
 ; US-09-872-136-4

Query Match 84.1%; Score 4429.5; DB 10; Length 993;
 Best Local Similarity 84.1%; Pred. No. 0;
 Matches 836; Conservative 57; Mismatches 90; Indels 11; Gaps 4;

QY 1 MRALAORSRRLLLVLSVMIETVNDLPVKVLI SHENNGSSAGKPSVYRVGRS 60
 DB 1 MPALA-RDAGTVPLLVVFSAMIFGTINQDLPVKVLIHNKNDSSVGKSSYPMVSES 59
 QY 61 PEDLQCTPRQSGTVVEAATVEAESGSIITLQVQLATPGDLSCLVWFKHSSLCQPHFD 120

Db 60 PEDLGCALPQSSGTYEAAVEVDVASITLQVLDPAGNISCLWFKHSSLNCPQHPD 119
Qy 121 LQNRGIVSMALNVETQAGEYLLHQSERANTYVLTNNVDTQTLVLRPFRKXENQ 180
Db 120 LQNRGVSVMLKWTQAGEYLLFTQSEATNTILFTVSRNTLLYTLRRPFRKXENQ 179
Qy 181 DALLCISEGPEPTVBLWLCSSHRECKEKGPAVVRKEEVLHELFGDTRCCARNALGR 240
Db 180 DALVCISESPEPIVEVWLCDQSGESCKESPAPVVRKEEVLHELFGDTRCCARNALGR 239
Qy 241 ECTKLETLNQAQOSTLPOLFLKVGPELWIRCKATHVNHGFLTWELDKALEEGSYFE 300
Db 240 ECTRLFTIDLNPQTLPOLFLKVGPELWIRCKATHVNHGFLTWELDKALEEGSYFE 299
Qy 301 MSTYSNTRMIRILLAFVSVSGRNDGYTCSSSKHPSQSAVLTILEKGFINATSSOEY 360
Db 300 MSTYSNTRMIRILFAFVSVVANDGTGYTCSSSKHPSQSAVLTIVGKGFINATSSSEY 359
Qy 361 EIDYKFCFVSFRKAYPRIRCTWISQASFPCEQRLGEGDGYISKPCDHKNKPGYIFY 420
Db 360 EIDQYEEFCFVSFRKAYPQIRCTWTSRKSFPCEQRLGEGDGYISKPCDHKNKPGYIFY 419
Qy 421 AENDDAQFTKMTNIRKPKQVLANASASQSCSDGYPLPSMTWKCKSKDNCTEIP 480
Db 420 AENDDAQFTKMTNIRKPKQVLAERASASQSCSDGYPLPSMTWKCKSKDNCTEIP 479
Qy 481 EGWNKANKRVFGQVWSSSTLMSBAGKLLVKCCAYNSMGTSCTIFLNSPGPPFIQ 540
Db 480 EGWNKANKRVFGQVWSSSTLMSBAGKLLVKCCAYNSMGTSCTIFLNSPGPPFIQ 539
Qy 541 DNISFYATIGLCLPFVILVILCHYKQKQFYESQLOMTQVGLDNEFYVDFRDY 600
Db 540 DNISFYATIGLCLPFVILVILCHYKQKQFYESQLOMTQVGLDNEFYVDFRDY 599
Qy 601 DLKWEPPRENLEFGKVLGSGAFGRVMTATYAGISKTGVSIOQAVKMLKEKADSCKEALM 660
Db 600 DLKWEPPRENLEFGKVLGSGAFGRVMTATYAGISKTGVSIOQAVKMLKEKADSCKEALM 659
Qy 661 SELKMTLGHGHDNIENLLGACTLSGPVYLI FEYCCYGDLLNLYRSKREKPHRTWTEIFK 720
Db 660 SELKMTLGHGHDNIENLLGACTLSGPVYLI FEYCCYGDLLNLYRSKREKPHRTWTEIFK 719
Qy 721 EHNFSYPTFOAHNSNMPGSRVQLHPDLSGFNGNSIHSEDETEYENQKRLABEE 780
Db 720 EHNFSYPTFOAHNSNMPGSRVQLHPDLSGFNGNSIHSEDETEYENQKRLABEE 777
Qy 781 EDNLVLTPEDLCLCFAYQVAKGMEFLFKSCVHRDLAARNVLTGHKVVVKI CDFGLARDIL 840
Db 778 EDNLVLTPEDLCLCFAYQVAKGMEFLFKSCVHRDLAARNVLTGHKVVVKI CDFGLARDIM 837
Qy 841 SDSYVVRGNARLPVKWMAPESEFEGYTTIKSDVMSYGILLWEIFSLGVNPPYGPV 900
Db 838 SDSYVVRGNARLPVKWMAPESEFEGYTTIKSDVMSYGILLWEIFSLGVNPPYGPV 897
Qy 901 FYKLIOGFMQOPFYATEGIIYFMQSCWAFDRKPSFNLTSFLGCQALAEAEAC 957
Db 898 FYKLIOGFMQOPFYATEGIIYFMQSCWAFDRKPSFNLTSFLGCQALAEAEAMQYN 957
Qy 958 ----IFTSIHLPKOAAAPQORG-GLRAQSPORQV 986
Db 958 VDRGVSECPHTYQNRPRFSEMDLGLLSPQAQVE 991

RESULT 5

US-10-192-867-2
; Sequence 2, Application US/10192867
; Publication No. US20030084466A1
; GENERAL INFORMATION:
; APPLICANT: BLUME-JENSEN, Peter
; APPLICANT: HUNTER, Tony
; TITLE OF INVENTION: SYSTEM AND METHOD FOR CONTROLLING MALE FERTILITY
; FILE REFERENCE: SALKINS.002C1

; CURRENT APPLICATION NUMBER: US/10/192,867
; CURRENT FILING DATE: 2002-07-08
; PRIOR APPLICATION NUMBER: US 60/175,625
; 2000-01-11
; PRIOR APPLICATION NUMBER: PCT/US01/00573
; 2001-01-08
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 975
; TYPE: PRT
; ORGANISM: MUS MUSCULUS
US-10-192-867-2

Query Match 23.2%; Score 1223; DB 15; Length 975;
Best Local Similarity 32.1%; Pred. No. 5.9e-94;
Matches 324; Conservative 172; Mismatches 331; Indels 182; Gaps 39;

Qy 36 CVLISHENNGSSAGKSSYRMVRGSPEDLOCTPRROSEGVYEAATVEVAESGSITLQVQ 95
Db 12 CVLLVLLRGQTATSPSA-----SPGEPPSIHPAQSELIVEAGD-TL--- 54
Qy 96 LATPGDLSCLE-----WVKHSSILGCPHF-----DLQNRGIVSMALNVTTETQAGEYLL 144
Db 55 -----SLTCIDPPFVWTFK-----TYENWENKKNEMIOEKA-----EATRTGYTC 98
Qy 145 HIOSERANYTVLFTVNRVD-TQYVLRPRPYFRKXENODALL-----CIS 187
Db 99 ---SNSNGLTSSYVVRDPKALFLVGLPLFGK-EDSDALVRCLPTDPQVSNYSLIECDG 154
Qy 188 EG-----VPEPTVEWLCSHRECKEKGPAVVRKEEVLHELFGDTRCCARNALGR 240
Db 155 KSLPTDLTFVPNP-----KAGITIKNVKGAYH-----RLCVRCAQR 191
Qy 241 ECTKLE-----PTIDLNOA-----POSTLPOL-----FLKVGPELWIRCKATHVNHGFLTW 286
Db 192 DGTWLSHDKETLKVREAIKAI PVVSVPTSHLLKKGDTFTVVCITIKDVSTSVNSMWLXON 251
Qy 287 -----ELBDKALEGSGYFEMSTYSTNRTMIRILLAFVSVSGRNDGYTCSSSKHPSQ 339
Db 252 PQPHIAQVKNHSHRGDF-----NYERQETLT-----ISSARVDDSGVFCYANNFTGS 301
Qy 340 SALVT---ILEKGFINATS-SQEEYEDIPYEKCFVSFRKAYPR-IRCTWIFSOASPPCE 394
Db 302 ANVTTLTKVVEKGFINISPVKNTVFTVDGENVDLVVEYBAYPKPEQOMIYMNRT--SA 359
Qy 395 ORGLE-----DGYISKPCDH-----KNKPGEYIFYAENDDAQFTKMTNIRKPKQV 443
Db 360 NKGKDYVKSNDKNIIRVYNQLRLTLKGTGGTYTFLVNSDASASVTFFNYVNTKPEIL 419
Qy 444 A--NASASQASCSGDGYPLPSWTWKCKSKDNCTEIP-PEGVWNKANKRVFGQVWSSS 500
Db 420 TYDLRLNGMLQCVABEFPPEPTIDWYFCTGABORCTTPVSPVDVQVQVNSVSPFKLVQVS 479
Qy 501 TLNMSBAGKLLVKCCAYNSMGTSCTIFLNSPGPPFIQDNISFYATIGLCLPFIV--- 557
Db 480 SIDSSVFRHNGTVECKASNDVGKS--SAFFN-----PAFKEIOIQAHTLFTPLLLIGFWVAAG 533
Qy 558 ---VLIVLICHYKQKQFYESQLOMI-QVTPGLDNEFYVDFRDYEDLWKEPFPRENLEF 613
Db 534 AMGIIVNLVLYKYLQKPMYEVQWKVVEEING---NNYVYIDPTLPDYHDKWEPFNRNLSF 590
Qy 614 KVLGSGAFGRVMTATYAGISKTGVSIOQAVKMLKEKADSCKEALMSLKMTHLGHHD 673
Db 591 GKTLLGAGAFKGVMTATYAGLIKSDAMTVAKMLKPSAHLITERALMSLKVLYLGNHM 650
Qy 674 NIVNLLGACTLSGPVYLI FEYCCYGDLLNLYRSKREKPF-----HRTWTEIFKEHNFSY 727
Db 651 NIVNLLGACTVGGTTLVITEYCCYGDLLNLYRSKREKPF-----HRTWTEIFKEHNFSY 710
Qy 728 PTOAHNSN---PGSREVLHPDLSGFNGNSIHSEDETEYENQKRLABEEEDLN 784
Db 711 PSCDS-SNEYMDMKPGVSYV-VPTKTDK-----RRSARIDSYIERDVTPTAIMEDDELALD 763


```
Db 109 VF---VRDPAKFLVDRSLYKGDNDTLVRCPLTDPVNTVNSLKGCKQKPLPKDLRFIPD 165
Qy 193 PTVEWVLCSSHRE-----SKKEGPAVVRKEKVLHFLPGTDIRCCARNALGRECTK 244
Db 166 PKAGIMIKSVRAYHRLCHLCHSVDOEG-----KSVLSEKFLKVRPAFK-----A 210
Qy 245 LFTIDLNOAPOSTLPQLFKVGEPLWIRCKAHVNHGFLTWELDEK--ALBEE----- 295
Db 211 VPVVSUSKA-----SYLLREGEERTVCTIKDVSSSVYTWKRENSOTKLQEKYNSWH 264
Qy 296 GSYEMSTYTNRTMIRILLAFVSSVGRNDGYTTCSSKHPQSALVTILE---KGFIN 352
Db 265 GDF-----NYERQATLT-----ISSARVNDSGVFCYANNTFGSANVTTLLEWDKGIN 314
Qy 353 A-TSSQBEYEDPYEKFCFSVRFKAYPR-IRCTWIFSOQSP--PCEQRGLEGYSISKFC 408
Db 315 IFPMINTVFNVDGENVDLIVEYEAFFPEHQOVIYMRFTFDKWEDEPKSENEINRIV 374
Qy 409 DHKN-----KPEYIFYAENDDAQFTKMTFLNIRKRPQVLA--NASASQASCSDDGYP 459
Db 375 SELHLTRLKGTGGTYTFLVNSDVNAALAFNVYNTKPEILTYDLVNGMLQCVAAAGFP 434
Qy 460 LPSWTWKCDKSPNCTEEI--PEGVWKKANRKFVQWVSSSTLNMSEAGKLLVKCCAY 518
Db 435 EPTIDWYFCPQTEQRCASVLPVDVQTLNNSGPPFGKLVQSSIDSSAFKHNGTVECKAY 494
Qy 519 NSMGTSCETIFLNSPGPPFFIQDN-----ISFYATIGLCLPFIIVLVLIC 564
Db 495 NDVGKT--SAYFN-----PAFGKNKEQIHPTLFTPLLIGFVIVAGM-----MCIIVMILT 544
Qy 565 HKYKQKRYESQLOMI-OVTGPDNEXEYVDFRDYEDLKWEPFRENLEFGKVLGSGAGF 623
Db 545 KYLQKPMYEVQWKVBEING---NNVYIIDPTQLPYDHKWEFFPNRLSFGKTLGAGAGF 601
Qy 624 RVNATAYGISTGVSQVAVKMLKEKADSEKALMSELKQMTLHGHNDINVLGACT 683
Db 602 KVEATAYGLIKSDAAMTVAVKMLKPSAHLTEREALMSELKVLVSLGNHMINVLGACT 661
Qy 684 LSGPVYLIFECYCGDILLNLYRSREKPHRTWTEIFKEHNFSSYPTFOAHNSNMPGSR 743
Db 662 IGGPTLVITEYCCYCGDILLNFRRLKRDSPICSKQEDHAEALYKYNLLHSEKSCSDSTNEY 721
Qy 744 VOLHPPLDOL---SGFNNGNSHSEDEIYENQKRLAEBEEDLNVLTLPEDLLCFAYQVAK 800
Db 722 MDMPKPGSVYVPTKADKRSVRIGSYIERDVTTPAIMEDEDELALDL---EDLLSFSYQVAK 778
Qy 801 GMEFLFKSCVHRDLAARNVLVTHGKVKICDFGLARDILSDSSYVVRGNARLPVKWMAP 860
Db 779 GMAFLASKNCIHRDLAARNILLTHGRITKICDFGLARDIKDSDNSYVVKGNARLPVKWMAP 838
Qy 861 ESLFEGYITKSDVWSGILLWEIFSLGVNYPGIPVDANTYKLIQSGFKMEQPPYATEG 920
Db 839 ESIFNCYVTFESDWSYGIFLWELFSLGSSYPGMPVDSPGVDSPKFKYKMKIEGFRMLSP 898
Qy 921 IYFVWQSWAFDSRKPSFPNLTSLFGLQALAE 953
Db 899 MYDIMKTCWDADPLKRPFTFKQIVOLIEKQISES 931
```

RESULT 8

```
US-10-192-867-4
; Sequence 4, Application US/10192867
; Publication No. US20030084466A1
; GENERAL INFORMATION:
; APPLICANT: BLUME-JENSEN, Peter
; APPLICANT: HUNTER, Tony
; TITLE OF INVENTION: SYSTEM AND METHOD FOR CONTROLLING MALE FERTILITY
; FILE REFERENCE: SALKINS.002C1
; CURRENT APPLICATION NUMBER: US/10/192,867
; PRIOR FILING DATE: 2002-07-08
; PRIOR APPLICATION NUMBER: US 60/175,625
; 2000-01-11
; PRIOR APPLICATION NUMBER: PCT/US01/00573
```

```
; 2001-01-08
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 976
; TYPE: PRT
; ORGANISM: HOMO SAPIENS
US-10-192-867-4
```

```
Query Match 22.7%; Score 1196; DB 15; Length 976;
Best Local Similarity 31.4%; Pred. No. 1.1e-91;
Matches 312; Conservative 172; Mismatches 361; Indels 148; Gaps 32;
```

```
Qy 36 CVLISHENNGSSAGKPPSSYRMVRGSPEDLQCTPRQSECTVVEAATVEVAESGSILOVQ 95
Db 12 CVLLLLLRVQTGSSOPS-----VSPGEPSPPSIHGKSDLIVRGD-EIRLL 57
Qy 96 LATPDLSCLWVFKHSSLCGPHFDLQNRGIVSMAILNVTQAGEYLLHIISERANYTV 155
Db 58 CTDPGFVK--WTFE---ILDETNEKNQEWITEKA---EATNTGKYTCTNKHGLNSIV 108
Qy 156 LFTVNVDR-TOLYVLRPRYPFRKMNQDALLCI-----SEG-----VPE 192
Db 109 VF---VRDPAKFLVDRSLYKGDNDTLVRCPLTDPVNTVNSLKGCKQKPLPKDLRFIPD 165
Qy 193 PTVEWVLCSSHRE-----SKKEGPAVVRKEKVLHFLPGTDIRCCARNALGRECTK 244
Db 166 PKAGIMIKSVRAYHRLCHLCHSVDOEG-----KSVLSEKFLKVRPAFK-----A 210
Qy 245 LFTIDLNOAPOSTLPQLFKVGEPLWIRCKAHVNHGFLTWELDEK--ALBEE----- 295
Db 211 VPVVSUSKA-----SYLLREGEERTVCTIKDVSSSVYTWKRENSOTKLQEKYNSWH 264
Qy 296 GSYEMSTYTNRTMIRILLAFVSSVGRNDGYTTCSSKHPQSALVTILE---KGFIN 352
Db 265 GDF-----NYERQATLT-----ISSARVNDSGVFCYANNTFGSANVTTLLEWDKGIN 314
Qy 353 A-TSSQBEYEDPYEKFCFSVRFKAYPR-IRCTWIFSOQSP--PCEQRGLEGYSISKFC 408
Db 315 IFPMINTVFNVDGENVDLIVEYEAFFPEHQOVIYMRFTFDKWEDEPKSENEINRIV 374
Qy 409 DHKN-----KPEYIFYAENDDAQFTKMTFLNIRKRPQVLA--NASASQASCSDDGYP 459
Db 375 SELHLTRLKGTGGTYTFLVNSDVNAALAFNVYNTKPEILTYDLVNGMLQCVAAAGFP 434
Qy 460 LPSWTWKCDKSPNCTEEI--PEGVWKKANRKFVQWVSSSTLNMSEAGKLLVKCCAY 518
Db 435 EPTIDWYFCPQTEQRCASVLPVDVQTLNNSGPPFGKLVQSSIDSSAFKHNGTVECKAY 494
Qy 519 NSMGTSCETIFLNSPGPPFFIQDN-----ISFYATIGLCLPFIIVLVLIC 564
Db 495 NDVGKT--SAYFN-----PAFGKNKEQIHPTLFTPLLIGFVIVAGM-----MCIIVMILT 544
Qy 565 HKYKQKRYESQLOMI-OVTGPDNEXEYVDFRDYEDLKWEPFRENLEFGKVLGSGAGF 623
Db 545 KYLQKPMYEVQWKVBEING---NNVYIIDPTQLPYDHKWEFFPNRLSFGKTLGAGAGF 601
Qy 624 RVNATAYGISTGVSQVAVKMLKEKADSEKALMSELKQMTLHGHNDINVLGACT 683
Db 602 KVEATAYGLIKSDAAMTVAVKMLKPSAHLTEREALMSELKVLVSLGNHMINVLGACT 661
Qy 684 LSGPVYLIFECYCGDILLNLYRSREKPHRTWTEIFKEHNFSSYPTFOAHNSNMPGSR 743
Db 662 IGGPTLVITEYCCYCGDILLNFRRLKRDSPICSKQEDHAEALYKYNLLHSEKSCSDSTNEY 721
Qy 744 VOLHPPLDOL---SGFNNGNSHSEDEIYENQKRLAEBEEDLNVLTLPEDLLCFAYQVAK 800
Db 722 MDMPKPGSVYVPTKADKRSVRIGSYIERDVTTPAIMEDEDELALDL---EDLLSFSYQVAK 778
Qy 801 GMEFLFKSCVHRDLAARNVLVTHGKVKICDFGLARDILSDSSYVVRGNARLPVKWMAP 860
Db 779 GMAFLASKNCIHRDLAARNILLTHGRITKICDFGLARDIKDSDNSYVVKGNARLPVKWMAP 838
```

QY 861 ESLFEGYITKSDVWSYGILLWEIFSLGVPYGPVDFANFYKLIQSGFMQEPFYVATEG 920
DB 839 ESIFNCVYTFSDVWSYGILLWEIFSLGSPYGPVDFANFYKLIQSGFMQEPFYVATEG 898
QY 921 IYFVMSQWAFDSRKRPSFNLTSPGLCOLAE 953
DB 899 MYDINKTCWADPLKRPYKQIVOLIEKQISES 931

RESULT 9
US-09-961-403-4
; Sequence 4, Application US/09961403
; Publication No. US20030077589A1
; GENERAL INFORMATION:
; APPLICANT: HE-STUMPF, HOLGER
; APPLICANT: HAENDLER, BERNARD
; APPLICANT: KRAETZSCHMAR, JOERN
; APPLICANT: KREFT, BERTHOLT
; APPLICANT: WINTERHAGER, ELKE
; APPLICANT: REGIDOR, PEDRO
; APPLICANT: SCOTTI, SIMONE
; TITLE OF INVENTION: METHOD FOR IN VITRO DIAGNOSIS OF ENDOMETRIOSIS
; CURRENT APPLICATION NUMBER: US/09/961,403
; CURRENT FILING DATE: 2001-09-25
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 4
; LENGTH: 1088
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-961-403-4

Query Match 21.9%; Score 1152.5; DB 11; Length 1088;
Best Local Similarity 32.1%; Pred. No. 6.1e-88;
Matches 314; Conservative 154; Mismatches 320; Indels 189; Gaps 33;

QY 94 VOLATPGDLSCLWVFKHSLGCGQPHFDLQNRGIVSMALNVTTQAGEY--LLHIQSERA 151
DB 39 VOLNSSFSLRC---FGESEVSNQ--YPMSEESSDVEIRN--EENNSGLPVTLEVSSASA 92

QY 152 NYTLFTVNVVTRDTQLYVLRPRYFRKMNQDALLCISEG-----VPEPTV-----EW 197
DB 93 AHTGLTYCTYNNHQ-----TEENE-----LEGRHIYIYVDPDPVAFVPLGMDTY 136

QY 198 VLCSHRES---CK---EEGPAVVRKEEVL-----HELFGTDIRCCARNALGR 240
DB 137 LIVIVEDDSAIIPCTTDPETVTLNSEGVPASVDSRQGNFTVGPYICEATVKGK 196

QY 241 ECTKLFTIDLNOAPOSTLPQLFL-----KVGEPLMIRCKAIHVNHGFLTW-----E 287
DB 197 ---KQTIIPFNVAALKATSELDLEMEALKTVYKSGETIVVTC-AVFNNEVVDLQWTYPE 252

QY 288 LEDKALEGSPYEMSTYTNRMIR-----ILLAPVSSVGR---NDTGYTTCSSSK--- 335
DB 253 VKGKI-----TMLBEIKVPSIKLYTLTVPEATVKDSDGYECAARQATR 297

QY 336 --HPSQSALVTILEKGF--INATSSQEEVEIDPYKFCFVRFKAYPRIRCTWIFSQASF 391
DB 298 EVKEMKKVTISVHEKGFIETKFTFQLE-AVNLHEVKGHVVEVRAYPPPRISLWKNLTL 356

QY 392 PCEQRGLDGYISIKFCD-----HNKK-----PGEYIFVAENDDAQFTKMTFLN 435
DB 357 -----IENLTEITDVEKIQIRYRSLKLRAKEEDSGHYTIVAQEDAVKSYTFELL 410

QY 436 IRKPKQVL-----ANASASQSCSDGYPLPSWTWKCKSDKSNCTEIEPEGVWNKA 488
DB 411 TQVFSSILDLVDHGGTGGTQVTRCTAGTPLPDIEMWICKD-IKKCNNETS---WTILA 466

QY 489 NRKVFQGWSS-----STLNMSSAGKLLVKKCAVNSMGTSCTIFLNSPG 534
DB 467 NN-----VSNIIITHSRDRSTVEGRVTFKVEETIAVRCLAKNLLGAENRELKVA--- 518

QY 535 PPFITQDNISFYATTGLCLPPIVWLVILVILCHYKKQFRYESOLOMIQVOTGLDNEYFYVD 594
DB 519 --PTLSELTVAAAVLLVLIIVIIISLIVWIKQKPRYIEIRWRVIESISPDGHEYIYVD 576

QY 595 FRDYEDLKWEFPRENLEFGKVLGSGAFGRVNVNATAYGISKTGVSIOQAVKMLKERADSC 654
DB 577 PMQLPYDSRWEPFRDGLVLRVLGSGAFGKVEGTAYGLSRQOPVMKVAVMKDKPTARSS 636

QY 655 EKEALMSLKMTHLGHHDNIIVNLLGACTLSGPVVLIFPYCCYGDLLNLRSLRREKXF--- 711
DB 637 EKQALMSLKNTHLGPLNIVNLLGACTKSGPIVITIEYCFYGDVLYVHLKRDSEFLSH 696

QY 712 -----HRTWTEIFKEHNFSSYPTFOAHNSNSMP--GSREVOHLHPP 749
DB 697 HPEKPKKELDIFGLNPADESTRSYVILSPENNGDYMMDKQADTTQVPMLEKEVSKYS 756

QY 750 LDQLSGFNGNSIHSDEIEYENQKLABEEEDLNVLTPEDLCLCFAYOVAKMEFLFKS 809
DB 757 IQRSLYDRPASYYKKSMLESDSEVNLLSDNNSGLTLL---DLSFTYQVARGMEFLASKN 813

QY 810 CVHRDLAARNVLVTHGKVVKICDFGLARDILSDSSYVVRGNARLPVKWMAPELSLFEGIYT 869
DB 814 CVHRDLAARNVLLAOGKIVKICDFGLARDIMHDSNVVSKGSTFLPVKWWAPESIFDNLYT 873

QY 870 IKSDVMSYGILLWEIFSLGVPYGPVDFANFYKLIQSGFMQEPFYVATEGYFVMSQW 929
DB 874 TILSDVMSYGILLWEIFSLGVPYGPVDMVDSFTYKIKSGYRMAKPDHATSEVYEIMVKW 933

QY 930 AFDSEKRPSPFNLTSPFL 946
DB 934 NSEPEKRPSPFHLSEIV 950

RESULT 10
US-09-769-987-2
; Sequence 2, Application US/09769987
; Patent No. US20020055129A1
; GENERAL INFORMATION:
; APPLICANT: Matsui, Toshimitsu
; APPLICANT: Aaronson, Stuart A.
; APPLICANT: Pierce, Jacalyn H.
; TITLE OF INVENTION: Antibodies for the Alpha Platelet-Derived Growth Factor Receptor
; FILE REFERENCE: 14014.026602
; CURRENT APPLICATION NUMBER: US/09/769,987
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 08/460,656
; PRIOR FILING DATE: 1995-06-02
; PRIOR APPLICATION NUMBER: US 08/439,095
; PRIOR FILING DATE: 1995-05-11
; PRIOR APPLICATION NUMBER: US 07/915,884
; PRIOR FILING DATE: 1992-07-20
; PRIOR APPLICATION NUMBER: US 07/308,282
; PRIOR FILING DATE: 1989-02-09
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1089
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence; No. US20020055129A1 =
; OTHER INFORMATION: synthetic construct
US-09-769-987-2

Query Match 21.9%; Score 1152.5; DB 9; Length 1089;
Best Local Similarity 32.1%; Pred. No. 6.1e-88;
Matches 314; Conservative 154; Mismatches 320; Indels 189; Gaps 33;

QY 94 VOLATPGDLSCLWVFKHSLGCGQPHFDLQNRGIVSMALNVTTQAGEY--LLHIQSERA 151
DB 39 VOLNSSFSLRC---FGESEVSNQ--YPMSEESSDVEIRN--EENNSGLPVTLEVSSASA 92

QY 152 NYTLFTVNVVTRDTQLYVLRPRYFRKMNQDALLCISEG-----VPEPTV-----EW 197

TELEFAX: 206-682-6031
TELEX: 372836
INFORMATION FOR SEQ ID NO: 36:
SEQUENCE CHARACTERISTICS:
LENGTH: 1089 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ
US-09-955-363-36

Query Match	21.9%	Score 1150.5	DB 10	Length 1089
Best Local Similarity	32.2%	Pred. No. 9e-88		
Matches 313	Conservative 154	Mismatches 327	Indels 177	Gaps 33
QY	94	VQLATPGDLSCLWVFKHSSLCGCPHFDLQNRGIYSMAILNVTETQAGEY--LLHIQISERA 151		
Db	39	VQUNSSPSLRC---FGBSEVSWQ--YPMSEESSDVEIRN--EENNSGLFVTVLEVSASA 92		
QY	152	NYTVLFVFNVRDQLYVLRPRYFRKMNQDALLCISEG-----VPEPTV-----EW 197		
Db	93	AHTGLTYCYNHTQ-----TEENE-----LEGRHIYIVPDPDAFVPLGMDTY 136		
QY	198	VLCSHRES---CK---EBGPAVVRKEEKVL-----HELFCTDIRCCARNALGR 240		
Db	137	LVIVEDDDSAIIPCRITDPTPTVLHNSGVVPASYPASYSRQGFNGCTFTVGPVCEATVKGK 196		
QY	241	ECTKLTFTIDNQAPOSTLPQLFL-----KVGEPILWIRCKAIHVNHGFLTW---E 287		
Db	197	---KFOTIPFNVALKATSELDLMEALKTVYKSGETIVVTC-AVFNEVVVDLQWTYPGE 252		
QY	288	LEDKALEEGSVFEMSTVSTNRTMIRIILAFVSSVGR---NDTGYYTCCSSK-----HPSQ 339		
Db	253	VKGGI-----TILEIKVPSIKLVTLTVPEATVKDSDGDEYCAARQATREVKEMK 303		
QY	340	SALVTILEKGF--INATSSOEYIEDPYEKFCSVRPKAYPRJCTRWIFSQAFFPCBQRG 397		
Db	304	KVTISVHEKGFIEIKTFSSLE-AVNLEHVKHFFVEVRAIYPPRISWLNKLTU----- 356		
QY	398	LEDGYSISKPCD-----HNKK-----PCEYIFYAENDDAQFTKMTFLNIRKKPQ 441		
Db	357	IENLTEITTDVEKIQIRYRSKLLIRAKEEDSGHYTIVAQNEADAVKSYTFELLTQVPSS 416		
QY	442	VL-----ANASASQASCSSDGYPLPSWTWKCKSDKSPNCTSEIPEGVWNKANKRVKFG 494		
Db	417	ILDVDDHHGSGTGQVTRCTAEGTPLFDIEMWICKO-IKKCNNETS---WTILANN----- 468		
QY	495	QWVSS-----STLNMSEAGKLLVKCCAYNSMGFSCETIFLNSPGPFPFIQ 540		
Db	469	--VSNIIITHSRDRSTVEGRVTFAKVEETIAVRCLAKNLLGAENRELKVA-----PTLR 522		
QY	541	DNISFYATIGLCPFTIVLVILVILCHYKXQFRYESQLOMIQVTGPLDNEYFYVDFRDYEY 600		
Db	523	SELTAAAVLVLLVIVISLIVLVWIKQPRYEIRWRVIESIPDGEHYIYVDPQMLPY 582		
QY	601	DLKWEPPRENLEFGKVLGSGAFGRVNNATAYGHSKTGVSIOQAVKMLKEKAOSCEKALM 660		
Db	583	DSRWEFFRDGLVLRVLGSGAFGVVEGTAYGLSRQPPVKNVAKVXMKLPTARSSEKQALM 642		
QY	661	SELKMWTHLGHNDINVLIGACTLSGPVYLITYEYCCYGLLNVILRSKREKF----- 711		
Db	643	SELKIMTHLGPUNIVNLIGACTSKSPYIITYEYCFYGLVNLVHKNRDSFSLHHPEKPK 702		
QY	712	-----HRTWTEIFEKHNFFSYPTFAQHSNSMP--GSRVQLHPPLDQLSG 755		
Db	703	KELDIFGLNPADESTSYVILSFENNGDYMDMKQADTTQVPMLEKREYKVKYDIOQSLY 762		
QY	756	FNNGSIHSEDEIYENQKELABEEEEEDLVNLTFFEDLLCFAYQVAKGMEFLFKSCVHRDL 815		
Db	763	DRPASVYKXKMULDSEKNLLSDNSEGLTLL---DLLSFTYQVARGMEFLASKVCVHRDL 819		
QY	816	ARBNVLVTHGKVVKICDFGLARDLILSDSSYVVRGNARLVPKMWAPBSLPEGYITIKSDVW 875		

Db	820	AARNVLLAQGIKVICDFGLARDIMHDSNNYVSKGSTFLPVKQMAPESTPDNIYTLTSDVW	879
Qy	876	SYGILLWEISFLGVNPPGIPVDANFYKLIQSGFQEQPFYATEGIYFMQSCWAFDSRK	935
Db	880	SYGILLWEISLGGTPYGGMMVDSTFYNKIKSGYRMKXPDHATSEVYIMVRCWSEPEK	939
Qy	936	RPSPFNLTSL	946
Db	940	RPSPFHLSEIV	950

RESULT 14

```

RES011 14
US-09-866-510-10
; Sequence 10, Application US/09866510
; Patent No. US20020111304A1
; GENERAL INFORMATION:
; APPLICANT: KAZLAUSKAS, ANDRIUS
; APPLICANT: IKUNO, YASUSHI
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING FIBROTIC DISEASES
; FILE REFERENCE: ERW-104.01
; CURRENT APPLICATION NUMBER: US/09/866,510
; CURRENT FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: 60/250,747
; PRIOR FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: 60/289,103
; PRIOR FILING DATE: 2001-05-07
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 1089
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-866-510-10

```

[illegible]

QY	535	PPFPFIQDNISFIATIGLCLPFI	VIVVLI	VLIVLICHKYKQFYRESQLQMIQV	TGLDNEYFYVD	594
DB	519	--FTLRSELTVAAAVLVLLV	IIISLIV	VVIMKQPRYEIRWRVIESISPDG	HEYIYVD	576
QY	595	PRDYEYDLKNEFPRENLE	FPKVLG	SGAFGRVWNATAYGISKTGYSIQ	AVAKMLKEKADSC	654
DB	577	PMQLPYDSRWEPFRDGL	VLGVLG	SGAFGKVVVEGTAYGLSRSQPVMK	VAVKMLKPTARSS	636
QY	655	EKALMSELKOWTHLGH	HDNI	VNLLGACTLSGPVVLIFEYCYCGDLLN	LYLSKREKF---	711
DB	637	EKQALMSELKIMTHL	GHPLNI	VNLLGACTKSGPIIITEYCFYGDV	VNLYLHKQRDSFLSH	696
QY	712	-----HRTWTEIFK	ENFSSYP	TFQAHNSMSP--	SGREVQLHPP	749
DB	697	HPEKPKKELDIFGLN	PADESTR	SYVILLSFENNQDYMDMKQADTTQ	VYVPMLEKREKYSKD	756
QY	750	LDQLSGFNGNHSIH	SEDEI	ENOKRLAEEBEDLNVLTFEDLL	CFAYOVAKGMEFLPKS	809
DB	757	IQRSLYDRPASYK	KKMLD	SEVKNLLSDNNEGTLT---	DLSTYQVAKGMEFLASN	813
QY	810	CVHRDLAARNVLV	THGKV	KICDFGLARDIISDSYVVVRGNAR	LPKVMMAPESLPEGIYT	869
DB	814	CVHRDLAARNVL	LAQGI	KVICDFGLARDIMHDSNYVSKGST	TFLEPMKWAPESIFDNL	873
QY	870	IKSDVMSYGILL	WEIFS	LGVPNPDPGIPVDANFYKLIQSG	FKMEQPFYATEGIYFVWQSCW	929
DB	874	TLSDVMSYGILL	WEIFS	LGSTPYPGMVDSTFYNKIKSGYMAK	PDHATSEYVEIMVKCW	933
QY	930	AFDSRKEPSPENL	T	SFL	946	
DB	934	NSEPEKRPSPFV	HLSEIV	950		

RESULT 15

US-09-866-510-4

; Sequence 4, Application US/09866510

; Patent No. US20020111304A1

; GENERAL INFORMATION:

; APPLICANT: KAZLAUSKAS, ANDRIUS

; APPLICANT: IKUNO, YASUSHI

; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING FIBROTIC DISEASES

; FILE REFERENCE: ERM-104.01

; CURRENT APPLICATION NUMBER: US/09/866,510

; CURRENT FILING DATE: 2001-05-25

; PRIOR APPLICATION NUMBER: 60/250,747

; PRIOR FILING DATE: 2000-12-01

; PRIOR APPLICATION NUMBER: 60/289,103

; PRIOR FILING DATE: 2001-05-07

; NUMBER OF SEQ ID NOS: 33

; SOFTWARE: PatentIn ver. 2.1.

; SEQ ID NO 4

; LENGTH: 1089

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-866-510-4

Query Match 21.8%; Score 1148.5; DB 10; Length 1089;

Best Local Similarity 32.0%; Pred. No. 1.3e-87;

Matches 313; Conservative 155; Mismatches 320; Indels 189; Gaps 33

QY	94	VOLATPGDLSCLW	FKHSLG	COHPFDLQNRGIVSMALINVT	ETOAGEX--	LLHIQSEPA	151
DB	39	VOLNSSFSLRC---	FGSEVSWQ--	YPMSEESSDVEIRN--	EENNSGLFPTV	LVLEVSASA	92
QY	152	NYTVLFTVNRD	TQYVLR	RRPYFRKMNODALLCISEG-----	VPEPTV-----	EW	197
DB	93	AHTGLTYCYNHTQ-----	TEENE-----	LEGRHIYYVDPD	PAVFLPGMTDY	136	
QY	198	VLCSHRES----	CK---	EEGPAVVRKXKVL-----	HELFGTDIRCARNALGR	240	
DB	137	LVIVEDDDSAI	PCRTTDP	ETETVTLHNSG	VVPASVDSRQGFNGT	FTVGPICEATVKRG	196

Qy	241	ECTKLF	IDLNQAP	QSTLPOL	FL-----KVGBPLWIRCAIHVHGGFLTW----	E 281			
Db	197	---KFQ	TIPENVYAL	KATSELD	MEALKTVYKSGETIVVTC-AVFNNVVDLQMTYPGE	252			
Qy	288	LEBKALEEG	SFEMSTV	STNRKTMIR-----	ILAFVSSVGR-----NDTGYTCSSSK----	335			
Db	253	VKGKI	-----TWLEE	KVPSIKLVYTLT	VEATVKDSGDVECAARQATR	297			
Qy	336	--HPSQ	SAVTLILEKGF--	INATSSOBEY	IDPYEKFCFSVRFKAYPIR	CTWTFPSQASF 391			
Db	298	EVKEMK	VVTISVHEKGF	IEIKPTFSOLE-AVN	HLVEVKHFVVEVAYPPPRIS	WLNKQNLTL 356			
Qy	392	PCBQ	RGLEDGYSISK	FCD-----HKNK-----	PGEYIFAENDDAQFTKMF	TLN 435			
Db	357	-----IEN	TEITTDVEK	TOEIRYRSKULIR	AKEEDSGHYTTVAQNEADV	KSTFELL 410			
Qy	436	IRKKPQVL--	ANASASQAS	CSGSDGYPLPSWT	WKCKSDKSPNCTEETEE	PEGVWNKKA 488			
Db	411	TQVPSS	ILDVDHGGSTG	GTVRCTAEGTLP	DIENMI CKD-IKKCN	NETS--WTLA 466			
Qy	489	NRKVFG	QWVSS-----	STLNMSEAG	KLLLVKCCAYNSMGT	SETTFLNSPG 534			
Db	467	NN-----	VSNII	TEIHSRDRST	VEGRVTFAKVBETIAVR	CLAKNLGAENRELK	LA--518		
Qy	535	PPPI	QDNI	SFYATIGCL	PIVVLVILICHYKKQ	PRYESOLOMIQVTPGL	DNEYFVVD 594		
Db	519	--PTLR	SELVAAAVLL	VILVILISLIV	VVIMWKQPRYEIR	RWVIESISPDG	HEYIYVD 576		
Qy	595	PRDYE	DLKWEFP	RENLEFGK	VLGSGAGFRVMNAT	YGISKTGVSIO	VAVKMLKEKADSC 654		
Db	577	PMQLPY	DRWKFP	PDGLVLR	GLGSGAGFKVVEGT	AYGLSSQPVMKVAV	NMLKPTARSS 636		
Qy	655	EKEAL	MSELK	MTHLGH	HDNIUNLLGACTLS	GPVYLFEYCYCY	GDLNLYLSKREKF---711		
Db	637	EKOAL	MSELK	IMTHLGH	PLNIVNLLGACTK	SGPTYIITEYCFY	GDLVNLHKNRDSFLSH 696		
Qy	712	-----HRT	WTEIFKE	HNFSYP	PTFOAHSSMB-	-GSRVEQLHPP	749		
Db	697	HPEKPK	KELDIFGL	NPADESTR	SVILISFENNGDY	MDMKQADTTQYV	PMLERKEVSKYSD 756		
Qy	750	LDQLS	GFGNGSIH	SEDEIEYEN	KQLAEBEEDNL	VLTFDLLCFAYO	AVAKMEFLEPKS 809		
Db	757	IQRSLY	DRPAS	YKKKSM	LDSEVKNLLSD	NSEGLTLL--DLLS	FTYQVARGMEFLASKN 813		
Qy	810	CVRH	DLAARNV	LTHGK	VVKIKCDGLARD	ILSDSSYVVRGNAR	LPUKVMAPESL	FEGLYT 869	
Db	814	CVRH	DLAARNV	LLAGQ	KIVIKCDGLARD	IMHDSNYVSKGST	FLPUKVMAPES	IFDNLYT 873	
Qy	870	IKSD	WVSGY	GILLWE	IFSLGVNVP	PGIIPVDANFY	KLIQSGFKMGE	OPFYATEGIFYV	NQSCW 929
Db	874	TLSD	VNSY	GILLWE	IFSLGGT	IFYGMVVD	STFYFNKIKSGY	RMKPDHATSE	VEYIEMVKCW 933
Qy	930	AFDR	KRKR	PSPNL	STFL 946				
Db	934	NSE	PEK	RP	SFYHL	SEIV 950			

Search completed: August 26, 2003, 07:41:07
Job time : 49.4771 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2003, 07:21:59 ; Search time 21.4892 Seconds
(without alignments)
1953.187 Million cell updates/sec

Title: US-09-919-408A-2

Perfect score: 5264

Sequence: 1 MRALAQRSDRLLLLVLSV.....RGGLRAQSPQRQVRIHRRS 992

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.*

- 1: /cgn2_6/ptodata/1/iaa/5A.COMB.pap.*
- 2: /cgn2_6/ptodata/1/iaa/5B.COMB.pap.*
- 3: /cgn2_6/ptodata/1/iaa/6A.COMB.pap.*
- 4: /cgn2_6/ptodata/1/iaa/6B.COMB.pap.*
- 5: /cgn2_6/ptodata/1/iaa/PCTUS.COMB.pap.*
- 6: /cgn2_6/ptodata/1/iaa/backfiles1.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5264	100.0	992	1 US-07-813-593-2	Sequence 2, Appli
2	5264	100.0	992	1 US-07-977-451-2	Sequence 2, Appli
3	5264	100.0	992	1 US-07-946-507-2	Sequence 2, Appli
4	5264	100.0	992	1 US-08-252-517-2	Sequence 2, Appli
5	5264	100.0	992	1 US-07-906-397A-2	Sequence 2, Appli
6	5264	100.0	992	1 US-08-601-891-2	Sequence 2, Appli
7	5264	100.0	992	2 US-09-021-324-2	Sequence 2, Appli
8	5264	100.0	992	5 PCT-US92-02750-2	Sequence 2, Appli
9	5264	100.0	992	5 PCT-US92-05401-2	Sequence 2, Appli
10	5264	100.0	992	5 PCT-US92-09893-2	Sequence 2, Appli
11	5102	96.9	1000	1 US-08-222-299-2	Sequence 2, Appli
12	5102	96.9	1000	1 US-08-434-878-2	Sequence 2, Appli
13	5102	96.9	1000	5 PCT-US95-03718-2	Sequence 2, Appli
14	4436.5	84.3	993	1 US-08-222-299-4	Sequence 4, Appli
15	4436.5	84.3	993	2 US-08-434-878-4	Sequence 4, Appli
16	4436.5	84.3	993	5 PCT-US95-03718-4	Sequence 4, Appli
17	4430.5	84.2	1160	5 PCT-US92-05401-4	Sequence 4, Appli
18	4429.5	84.1	993	1 US-07-977-451-4	Sequence 4, Appli
19	4429.5	84.1	993	1 US-08-252-517-4	Sequence 4, Appli
20	4429.5	84.1	993	1 US-07-906-397A-4	Sequence 4, Appli
21	4429.5	84.1	993	1 US-08-601-891-4	Sequence 4, Appli
22	4429.5	84.1	993	2 US-08-021-324-4	Sequence 4, Appli
23	4429.5	84.1	993	5 PCT-US92-09893-4	Sequence 4, Appli
24	4408.5	83.7	993	1 US-08-183-211-2	Sequence 2, Appli
25	4408.5	83.7	993	5 PCT-US95-00176A-2	Sequence 2, Appli
26	2240.5	42.6	481	3 US-07-912-122-4	Sequence 4, Appli
27	2240.5	42.6	481	5 PCT-US93-06404-4	Sequence 4, Appli

Query Match 100.0%; Score 5264; DB 1; Length 992;

28 1221.5 23.2 972 3 US-08-750-141A-2 Sequence 2, Appli
29 1196 22.7 976 3 US-08-750-141A-1 Sequence 1, Appli
30 1152.5 21.9 1089 1 US-08-168-917-4 Sequence 4, Appli
31 1152.5 21.9 1089 2 US-08-460-510-4 Sequence 4, Appli
32 1152.5 21.9 1089 2 US-08-460-490-4 Sequence 4, Appli
33 1152.5 21.9 1089 3 US-08-462-728-2 Sequence 2, Appli
34 1152.5 21.9 1089 4 US-08-461-917-2 Sequence 2, Appli
35 1152.5 21.9 1089 4 US-08-464-436-2 Sequence 2, Appli
36 1152.5 21.9 1089 5 PCT-US92-00730-4 Sequence 4, Appli
37 1152.5 21.9 1089 5 PCT-US92-00862-4 Sequence 4, Appli
38 1150.5 21.9 1089 1 US-08-180-129-36 Sequence 36, Appli
39 1150.5 21.9 1089 1 US-08-477-329-36 Sequence 36, Appli
40 1150.5 21.9 1089 2 US-08-475-458-36 Sequence 36, Appli
41 1150.5 21.9 1089 3 US-08-980-400-36 Sequence 36, Appli
42 1150.5 21.9 1089 3 US-09-583-459A-36 Sequence 36, Appli
43 1150.5 21.9 1089 3 US-09-583-210-36 Sequence 36, Appli
44 1150.5 21.9 1089 4 US-09-583-449A-36 Sequence 36, Appli
45 1150.5 21.9 1089 4 US-09-435-059-36 Sequence 36, Appli

ALIGNMENTS

RESULT 1

US-07-813-593-2
; Sequence 2, Application US/07813593
; Patent No. 5185438
; GENERAL INFORMATION:
; APPLICANT: Lemischka, Ihor R.
; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL
; TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: IMCLONE SYSTEMS INCORPORATED
; STREET: 180 VARICK STREET
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: U.S.A.
; ZIP: 10014
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/813,593
; FILING DATE: 19920415
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/793,065
; FILING DATE: 15-NOV-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/728,913
; FILING DATE: 28-JUN-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/679,666
; FILING DATE: 02-APR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Feit, Irving N.
; REGISTRATION NUMBER: 28,601
; REFERENCE/DOCKET NUMBER: LEM-3-PPP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-645-1405
; TELEFAX: 212-645-2054
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 992 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-813-593-2

Best Local Similarity 100.0%; Pred. No. 0;		Matches 992; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1	MRALQSRDRRLLLVLSVMIETVTNODLPVKVCLISHENNGSSACKPSSYRMVRS	60
Db	1	MRALQSRDRRLLLVLSVMIETVTNODLPVKVCLISHENNGSSACKPSSYRMVRS	60
QY	61	PEDLOCTPRQSEGTYVEATVEABSGSITLQVQLATPGDLSCLVVFKHSSIGCQPHD	120
Db	61	PEDLOCTPRQSEGTYVEATVEABSGSITLQVQLATPGDLSCLVVFKHSSIGCQPHD	120
QY	121	LQNRGIVSMALNVETQAGEYLLHIQSERANYTVLFTVNVNRTDQLYVLRPPYFRKMEQ	180
Db	121	LQNRGIVSMALNVETQAGEYLLHIQSERANYTVLFTVNVNRTDQLYVLRPPYFRKMEQ	180
QY	181	DALLCISEGVPTVEWVLCSSHRESCKEEGPAVVRKEKVLHELPGTDIRCARNALGR	240
Db	181	DALLCISEGVPTVEWVLCSSHRESCKEEGPAVVRKEKVLHELPGTDIRCARNALGR	240
QY	241	ECTKLTIDLNQAPOSTLPQLFLKVGEPWIRCKAIHVNHGFGLTWELEDKALEGSYPE	300
Db	241	ECTKLTIDLNQAPOSTLPQLFLKVGEPWIRCKAIHVNHGFGLTWELEDKALEGSYPE	300
QY	301	MSTYSTNRTMIRILLAFVSSVGRNDTGYTCSSSKHPSQSALVTILEKGFINATSQEY	360
Db	301	MSTYSTNRTMIRILLAFVSSVGRNDTGYTCSSSKHPSQSALVTILEKGFINATSQEY	360
QY	361	EIDPYEKECFVSFRKAYPRICWTWFSQAFCEQRLGDEGYSISKPCDHKNKPGYIEY	420
Db	361	EIDPYEKECFVSFRKAYPRICWTWFSQAFCEQRLGDEGYSISKPCDHKNKPGYIEY	420
QY	421	AENDDAQFTKMTLIRKPPQVLNANASQASQSSDGYPLPMTWKKCDKSPNCTEELP	480
Db	421	AENDDAQFTKMTLIRKPPQVLNANASQASQSSDGYPLPMTWKKCDKSPNCTEELP	480
QY	481	EGVWNNKANRVFGQWSSSTLNMSAGKGLLVKCCAYNSMGTSCETIFLNSPGPPFTQ	540
Db	481	EGVWNNKANRVFGQWSSSTLNMSAGKGLLVKCCAYNSMGTSCETIFLNSPGPPFTQ	540
QY	541	DNISFYATIGLCPFTVLVLI CHYKQFYESOLOMIQVTPGLDNEYFYVDFRDVEY	600
Db	541	DNISFYATIGLCPFTVLVLI CHYKQFYESOLOMIQVTPGLDNEYFYVDFRDVEY	600
QY	601	DLKWEFFPRENLFFGVLGSGAFGRVNNATAYGTSKTVSIQVAVKMLKEKADSCKEALM	660
Db	601	DLKWEFFPRENLFFGVLGSGAFGRVNNATAYGTSKTVSIQVAVKMLKEKADSCKEALM	660
QY	661	SELKMWTHLGHNDINVLGACTLSGPVYLI FEYCCYGDLLNLYRSKREKFKHRTWTIPK	720
Db	661	SELKMWTHLGHNDINVLGACTLSGPVYLI FEYCCYGDLLNLYRSKREKFKHRTWTIPK	720
QY	721	EHNFSSYPTFOAHSSNMPGSRVQLHPDLQSGFNNGSIHSEDEIEYENOKRLAEBEE	780
Db	721	EHNFSSYPTFOAHSSNMPGSRVQLHPDLQSGFNNGSIHSEDEIEYENOKRLAEBEE	780
QY	781	EDNLVLTEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVKICDFGLARDIL	840
Db	781	EDNLVLTEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVKICDFGLARDIL	840
QY	841	SDSSYVVRGNARLPVKWMAPELFGIYTIKSDVMSYGILLWEIFSLGWNYPYGPVDAN	900
Db	841	SDSSYVVRGNARLPVKWMAPELFGIYTIKSDVMSYGILLWEIFSLGWNYPYGPVDAN	900
QY	901	FYKLIQSGFKMEQPYATYEGIFVWQSCWAFPSRKRPSFPNLTSLGCOLAEBACIRT	960
Db	901	FYKLIQSGFKMEQPYATYEGIFVWQSCWAFPSRKRPSFPNLTSLGCOLAEBACIRT	960
QY	961	SIHLPKQAAPQORGGLRAQSPORQVKIHRERS 992	
Db	961	SIHLPKQAAPQORGGLRAQSPORQVKIHRERS 992	

US-07-977-451-2
; Sequence 2, Application US/07977451
; Patent No. 5270458
; GENERAL INFORMATION:
; APPLICANT: Lemischka, Ihor R.
; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL
; TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ImClone Systems Incorporated
; STREET: 180 Varick Street
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10014
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/977,451
; FILING DATE: 19921119
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US UNASSIGNED
; FILING DATE: 12-NOV-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/906,397
; FILING DATE: 26-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US PCT/US92/05401
; FILING DATE: 26-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: TW 81102961
; FILING DATE: 15-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US PCT/US92/02750
; FILING DATE: 02-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/813,593
; FILING DATE: 24-DEC-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/793,065
; FILING DATE: 15-NOV-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/728,913
; FILING DATE: 28-JUN-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/679,666
; FILING DATE: 02-APR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Feit, Irving N.
; REGISTRATION NUMBER: 28,601
; REFERENCE/DOCKET NUMBER: LEM-3-7P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-645-1405
; TELEFAX: 212-645-2054
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 992 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-977-451-2

Query Match		100.0%; Score 5264; DB 1; Length 992;	
Best Local Similarity		100.0%; Pred. No. 0;	
Matches 992; Conservative		0; Mismatches 0; Indels 0; Gaps 0;	
QY	1	MRALQSRDRRLLLVLSVMIETVTNODLPVKVCLISHENNGSSACKPSSYRMVRS	60
Db	1	MRALQSRDRRLLLVLSVMIETVTNODLPVKVCLISHENNGSSACKPSSYRMVRS	60

Qy 61 PEDLOCTPRROSEGTVEAATVEAEGSSITLQVQLATPGDLSCLVFVFKHSSIGCQPHFD 120
Db 61 PEDLOCTPRROSEGTVEAATVEAEGSSITLQVQLATPGDLSCLVFVFKHSSIGCQPHFD 120
Qy 121 LQNRGIVSMALNVTTQAGEYLLHIQSERANYTLFTVNVVDTQLYVLRPPYFRKMNQ 180
Db 121 LQNRGIVSMALNVTTQAGEYLLHIQSERANYTLFTVNVVDTQLYVLRPPYFRKMNQ 180
Qy 181 DALLCISGEVPEPTVEWVLCSSHRESCKEGPAVVRKEEVHLFGTDIRCCARNALGR 240
Db 181 DALLCISGEVPEPTVEWVLCSSHRESCKEGPAVVRKEEVHLFGTDIRCCARNALGR 240
Qy 241 ECTKLFITDLNQAPQSTLPQLFKVGEPLIRCKAIHVNHGFGLTWELEDKALEGGSYFE 300
Db 241 ECTKLFITDLNQAPQSTLPQLFKVGEPLIRCKAIHVNHGFGLTWELEDKALEGGSYFE 300
Qy 301 MSTYSTNTRTIRILLAFVSSVGRNDTGYTTCSSSKHPSQSALVTILEKGFINATSSQBEY 360
Db 301 MSTYSTNTRTIRILLAFVSSVGRNDTGYTTCSSSKHPSQSALVTILEKGFINATSSQBEY 360
Qy 361 BIDPYEKFCSFVRKAYPRIRCTWIFSOAEPCEQGLEGDSYISKFCDHKNKPGYIFY 420
Db 361 BIDPYEKFCSFVRKAYPRIRCTWIFSOAEPCEQGLEGDSYISKFCDHKNKPGYIFY 420
Qy 421 AENDDAQFTKFTNIRKPKQVLNASASQASCSDDGYPLSWTWKCKSDKSPNCTEIP 480
Db 421 AENDDAQFTKFTNIRKPKQVLNASASQASCSDDGYPLSWTWKCKSDKSPNCTEIP 480
Qy 481 EGVNKKANRVFGOWSSSTLNSEAGKLLVKCCAYNSMGTSCETIFLNSPGPFPIQ 540
Db 481 EGVNKKANRVFGOWSSSTLNSEAGKLLVKCCAYNSMGTSCETIFLNSPGPFPIQ 540
Qy 541 DNISFYATIGLCPFIVVLVILCHIKYKQFRIESQLQMIQVTPGLDNEYFYVDFRDY 600
Db 541 DNISFYATIGLCPFIVVLVILCHIKYKQFRIESQLQMIQVTPGLDNEYFYVDFRDY 600
Qy 601 DLKWEFPRENLEFGKVLGSGAFGRVMNATAYGISTGVSIQVAVQMLKEKADSCKEALM 660
Db 601 DLKWEFPRENLEFGKVLGSGAFGRVMNATAYGISTGVSIQVAVQMLKEKADSCKEALM 660
Qy 661 SELQWTHLGHNDIVNLLGACTLSGPVYLIFCYCCYDGLNLYRSKEKEHRTWTETPK 720
Db 661 SELQWTHLGHNDIVNLLGACTLSGPVYLIFCYCCYDGLNLYRSKEKEHRTWTETPK 720
Qy 721 EHNFSYPTFOAHNSNMPGSRVQLHPPLDQLSGFNNGSIHSEDEIEYENQKRLAEBEE 780
Db 721 EHNFSYPTFOAHNSNMPGSRVQLHPPLDQLSGFNNGSIHSEDEIEYENQKRLAEBEE 780
Qy 781 EDNLVLTFFEDLLCFAYQVAKWMEFLFKSCVHRDLAARNVLVTHGKVVKICDFGLARDIL 840
Db 781 EDNLVLTFFEDLLCFAYQVAKWMEFLFKSCVHRDLAARNVLVTHGKVVKICDFGLARDIL 840
Qy 841 SDSYVVRGNARLPVKWMAPELSEFEGYITIKSDVMSYGLLWEIFSLGVNYPGIPVDAN 900
Db 841 SDSYVVRGNARLPVKWMAPELSEFEGYITIKSDVMSYGLLWEIFSLGVNYPGIPVDAN 900
Qy 901 FYKLIQSGFKMEQPFYATEGIFYFVWQSCWAFDSKRPSFPNLTSLFGQLAEBEACIRT 960
Db 901 FYKLIQSGFKMEQPFYATEGIFYFVWQSCWAFDSKRPSFPNLTSLFGQLAEBEACIRT 960
Qy 961 SIHLPKQAAPQOQGLRAQSPQVKIHRNS 992
Db 961 SIHLPKQAAPQOQGLRAQSPQVKIHRNS 992

RESULT 3

US-07-946-507-2
; Sequence 2, Application US/07946507
; Patent No. 5283354
; GENERAL INFORMATION:
; APPLICANT: Lemischka, Ihor R.
; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL

; TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: IMCLONE SYSTEMS INCORPORATED
; STREET: 180 VARICK STREET
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: U.S.A.
; ZIP: 10014
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/946,507
; FILING DATE: 19920917
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/813,593
; FILING DATE: 24-DEC-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/793,065
; FILING DATE: 15-NOV-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/728,913
; FILING DATE: 28-JUN-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/679,666
; FILING DATE: 02-APR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Feit, Irving N.
; REGISTRATION NUMBER: 28,601
; REFERENCE/DOCKET NUMBER: LEM-3-PPP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-645-1405
; TELEFAX: 212-645-2054
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 992 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-946-507-2

Query Match 100.0%; Score 5264; DB 1; Length 992;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 992; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MRALQASDRRLLLVLSVMILETVTNQDLPVIKCVLISHENNGSSAGKPSYSYRWGRS 60
Db 1 MRALQASDRRLLLVLSVMILETVTNQDLPVIKCVLISHENNGSSAGKPSYSYRWGRS 60
Qy 61 PEDLOCTPRROSEGTVEAATVEAEGSSITLQVQLATPGDLSCLVFVFKHSSIGCQPHFD 120
Db 61 PEDLOCTPRROSEGTVEAATVEAEGSSITLQVQLATPGDLSCLVFVFKHSSIGCQPHFD 120
Qy 121 LQNRGIVSMALNVTTQAGEYLLHIQSERANYTLFTVNVVDTQLYVLRPPYFRKMNQ 180
Db 121 LQNRGIVSMALNVTTQAGEYLLHIQSERANYTLFTVNVVDTQLYVLRPPYFRKMNQ 180
Qy 181 DALLCISGEVPEPTVEWVLCSSHRESCKEGPAVVRKEEVHLFGTDIRCCARNALGR 240
Db 181 DALLCISGEVPEPTVEWVLCSSHRESCKEGPAVVRKEEVHLFGTDIRCCARNALGR 240
Qy 241 ECTKLFITDLNQAPQSTLPQLFKVGEPLIRCKAIHVNHGFGLTWELEDKALEGGSYFE 300
Db 241 ECTKLFITDLNQAPQSTLPQLFKVGEPLIRCKAIHVNHGFGLTWELEDKALEGGSYFE 300
Qy 301 MSTYSTNTRTIRILLAFVSSVGRNDTGYTTCSSSKHPSQSALVTILEKGFINATSSQBEY 360
Db 301 MSTYSTNTRTIRILLAFVSSVGRNDTGYTTCSSSKHPSQSALVTILEKGFINATSSQBEY 360

Qy 361 EIDPYEKFCSVRFKAYPRIRCTWIFSQAQSPFCEQGLGDSYISKFCDHKNKPGYIFY 420
Db 361 EIDPYEKFCSVRFKAYPRIRCTWIFSQAQSPFCEQGLGDSYISKFCDHKNKPGYIFY 420
Qy 421 AENDDAQTKMFTLNIRKPKQVLANASQASQSSDGYPLPSWTWKKCSDKSPNCTEIP 480
Db 421 AENDDAQTKMFTLNIRKPKQVLANASQASQSSDGYPLPSWTWKKCSDKSPNCTEIP 480
Qy 481 EGVNKKANRVFGQWSSSTLNMSGAGLLVKCCAYNSMGTSCETIFLNSPGFPFIQ 540
Db 481 EGVNKKANRVFGQWSSSTLNMSGAGLLVKCCAYNSMGTSCETIFLNSPGFPFIQ 540
Qy 541 DNISFYATIGLCPPIVVLIVLICHYKQFYESOLOMIQVTPLDNEYFYVDRDYEY 600
Db 541 DNISFYATIGLCPPIVVLIVLICHYKQFYESOLOMIQVTPLDNEYFYVDRDYEY 600
Qy 601 DLKWEFPRENLEFGKVLGSGAFGRVYNATAYGISKTVGSIQAVKMLKEKADSCKEALM 660
Db 601 DLKWEFPRENLEFGKVLGSGAFGRVYNATAYGISKTVGSIQAVKMLKEKADSCKEALM 660
Qy 661 SELKMTLHGHNDINVLGACTLGGPVYLIPEYCYGDLNLYRSKREKPHRTWTEIFK 720
Db 661 SELKMTLHGHNDINVLGACTLGGPVYLIPEYCYGDLNLYRSKREKPHRTWTEIFK 720
Qy 721 EHNFSYPTFOAHNSMSPGSRVQLHPPDLQSLGFGNSIHSEDEIEYENKRLAEBEE 780
Db 721 EHNFSYPTFOAHNSMSPGSRVQLHPPDLQSLGFGNSIHSEDEIEYENKRLAEBEE 780
Qy 781 EDNLVLTFFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKWKICDFGLARDIL 840
Db 781 EDNLVLTFFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKWKICDFGLARDIL 840
Qy 841 SDSSVVRGNARLPVKWMAPELSFEGYITIKSDVMSYGILLWEIFSLGWNVPYGPVVDAN 900
Db 841 SDSSVVRGNARLPVKWMAPELSFEGYITIKSDVMSYGILLWEIFSLGWNVPYGPVVDAN 900
Qy 901 FYKLIQSFGKMPQFYATYEGYFVVMQSWAFDRSRKPSFPNLTSLGCOLAEBEACIRT 960
Db 901 FYKLIQSFGKMPQFYATYEGYFVVMQSWAFDRSRKPSFPNLTSLGCOLAEBEACIRT 960
Qy 961 SIHLPKQAAPQORGGLRAQSPORQVKIHRERS 992
Db 961 SIHLPKQAAPQORGGLRAQSPORQVKIHRERS 992

RESULT 4
US-08-252-517-2
; Sequence 2, Application US/08252517
; Patent No. 5548065
; GENERAL INFORMATION:
; APPLICANT: Lemischka, Ihor R.
; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL
; TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ImClone Systems Incorporated
; STREET: 180 Varick Street
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10014
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/252,517
; FILING DATE: 31-OCT-1994
; CLASSIFICATION: 530
; -PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/977,451
; FILING DATE: 19-NOV-1992

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/906,397
; FILING DATE: 26-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US PCT/US92/05401
; FILING DATE: 26-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: TW 81102961
; FILING DATE: 15-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US PCT/US92/02750
; FILING DATE: 02-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/813,593
; FILING DATE: 24-DEC-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/793,065
; FILING DATE: 15-NOV-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/728,913
; FILING DATE: 28-JUN-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/679,666
; FILING DATE: 02-APR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Feit, Irving N.
; REGISTRATION NUMBER: 28,601
; REFERENCE/DOCKET NUMBER: LEM-3-7P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-645-1405
; TELEFAX: 212-645-2054
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 992 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-252-517-2

Query Match 100.0%; Score 5264; DB 1; Length 992;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 992; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MEALAQRSDRRLLLVLSVMILETTVNTQDLPVKCVLISHENNGSSACKPSYRMVRS 60
Db 1 MEALAQRSDRRLLLVLSVMILETTVNTQDLPVKCVLISHENNGSSACKPSYRMVRS 60
Qy 61 PEDLOCTPRQSEGTVEAATVEVAESGSIITLQVQLATPGDLSCLVWFKXSLGCOPHFD 120
Db 61 PEDLOCTPRQSEGTVEAATVEVAESGSIITLQVQLATPGDLSCLVWFKXSLGCOPHFD 120
Qy 121 LQNRGIVSMALNVTETQAGEYLLHIQSERANYTLVTVNVRDTOLYVLRPPYFRKXWQ 180
Db 121 LQNRGIVSMALNVTETQAGEYLLHIQSERANYTLVTVNVRDTOLYVLRPPYFRKXWQ 180
Qy 181 DALLCISGVPEPTVEWVLCSSHRESCKEEGPAVVRKEEKVLHFLGTDIRCCARNALGR 240
Db 181 DALLCISGVPEPTVEWVLCSSHRESCKEEGPAVVRKEEKVLHFLGTDIRCCARNALGR 240
Qy 241 ECTKLTIDLNQAPOSTLPOLFLKVGEPILWIRKAIHNVHNGGLTWELDKALEEGSYFE 300
Db 241 ECTKLTIDLNQAPOSTLPOLFLKVGEPILWIRKAIHNVHNGGLTWELDKALEEGSYFE 300
Qy 301 MSTYSTNRTMIRILLAFVSSVGRNDTGYVTCSSSKHPSOSALVTILEKGFINATSSQEEY 360
Db 301 MSTYSTNRTMIRILLAFVSSVGRNDTGYVTCSSSKHPSOSALVTILEKGFINATSSQEEY 360
Qy 361 EIDPYEKFCSVRFKAYPRIRCTWIFSQAQSPFCEQGLGDSYISKFCDHKNKPGYIFY 420
Db 361 EIDPYEKFCSVRFKAYPRIRCTWIFSQAQSPFCEQGLGDSYISKFCDHKNKPGYIFY 420
Qy 421 AENDDAQTKMFTLNIRKPKQVLANASQASQSSDGYPLPSWTWKKCSDKSPNCTEIP 480

Db 421 AENDDAQFTKMTLIRKKPOVLANASASQSDGYPLPSWTWKCKSDKSPNCTEIP 480
Qy 481 EGVNKKANRKYFGQWSSSTLNSEAGGLLVKCCAYNSMGTSCETIFLNSPGPPFIQ 540
Db 481 EGVNKKANRKYFGQWSSSTLNSEAGGLLVKCCAYNSMGTSCETIFLNSPGPPFIQ 540
Qy 541 DNISFYATIGLCPFIIVLVILCHIKYKQFRIESQLOMIQVTPGLDNEYFYVDFRDY 600
Db 541 DNISFYATIGLCPFIIVLVILCHIKYKQFRIESQLOMIQVTPGLDNEYFYVDFRDY 600
Qy 601 DLKWEFPRENLEFGKVLGSGAFGRVWNATAYIGSTGVSIQVAVKMLKEKADSCKEALM 660
Db 601 DLKWEFPRENLEFGKVLGSGAFGRVWNATAYIGSTGVSIQVAVKMLKEKADSCKEALM 660
Qy 661 SELKQWTHLGHHDNIIVNLGACTLSPVYLIFEYCCYDGLNLYLSKREKPHRTWTEIPK 720
Db 661 SELKQWTHLGHHDNIIVNLGACTLSPVYLIFEYCCYDGLNLYLSKREKPHRTWTEIPK 720
Qy 721 EHNFSYPTFOAHSSNMPGSRVQLHPPLDQSGFNNGSIHSEDEIEYENOKRLAEBEE 780
Db 721 EHNFSYPTFOAHSSNMPGSRVQLHPPLDQSGFNNGSIHSEDEIEYENOKRLAEBEE 780
Qy 781 EDLNVLTPEDDLFCFAYQVAKMEFLFKSCVHRDLAARNVLVTHGKVKICDFGLARDIL 840
Db 781 EDLNVLTPEDDLFCFAYQVAKMEFLFKSCVHRDLAARNVLVTHGKVKICDFGLARDIL 840
Qy 841 SDSSVVRGNARLPVKWMAPELSPFEGIYTIKSDVMSYGILLWEIFSLGVNPPGIPVDAN 900
Db 841 SDSSVVRGNARLPVKWMAPELSPFEGIYTIKSDVMSYGILLWEIFSLGVNPPGIPVDAN 900
Qy 901 FYKLIQSGFKMEOPPYATEGIFVQMSQWAFDRKRPSFPNLTSLGCOLABAEACIRT 960
Db 901 FYKLIQSGFKMEOPPYATEGIFVQMSQWAFDRKRPSFPNLTSLGCOLABAEACIRT 960
Qy 961 SIHLPKQAAPQOGRGLRAQSPQKVIHRS 992
Db 961 SIHLPKQAAPQOGRGLRAQSPQKVIHRS 992

RESULT 5
US-07-906-397A-2
; Sequence 2, Application US/07906397A
; Patent No. 5621090
; GENERAL INFORMATION:
; APPLICANT: Lemischka, Ihor R.
; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL
; TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCLONE SYSTEMS INCORPORATED
; STREET: 180 VARICK STREET
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: U.S.A.
; ZIP: 10014
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/906,397A
; FILING DATE: 19920626
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/813,593
; FILING DATE: 24-DEC-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/793,065
; FILING DATE: 15-NOV-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/728,913
; FILING DATE: 28-JUN-1991

;; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/679,666
; FILING DATE: 02-APR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Feit, Irving N.
; REGISTRATION NUMBER: 28,601
; REFERENCE/DOCKET NUMBER: LEM-3-PPPPPP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-645-1405
; TELEFAX: 212-645-2054
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 992 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-906-397A-2

Query Match 100.0%; Score 5264; DB 1; Length 992;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 992; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MRALAQRSDRRLLLVLSVMIETVTNODLPVICKVLISHENNGSSAGKPPSSYRMVRGS 60
Db 1 MRALAQRSDRRLLLVLSVMIETVTNODLPVICKVLISHENNGSSAGKPPSSYRMVRGS 60
Qy 61 PEDLQCTPRRQSEGTVEAATVEVAESGSIITLQVQLATPGDLSCLVWFKHSSILGCOPHFD 120
Db 61 PEDLQCTPRRQSEGTVEAATVEVAESGSIITLQVQLATPGDLSCLVWFKHSSILGCOPHFD 120
Qy 121 LQNRGIVSMALINVTETQAGEYLLHIQSERANYTVLFTVNVNRTDQLYLRRPYFRKMEQ 180
Db 121 LQNRGIVSMALINVTETQAGEYLLHIQSERANYTVLFTVNVNRTDQLYLRRPYFRKMEQ 180
Qy 181 DALLCISGVPEPTVEWVLCSSHRESCKEEGPAVVRKEEKLHELFGTDIRCCARNALGR 240
Db 181 DALLCISGVPEPTVEWVLCSSHRESCKEEGPAVVRKEEKLHELFGTDIRCCARNALGR 240
Qy 241 ECTKLFTIDLNQAPQSTLPQLFKVGEPLWIRCKAIHVNHGFLTWELDKALEGSYFE 300
Db 241 ECTKLFTIDLNQAPQSTLPQLFKVGEPLWIRCKAIHVNHGFLTWELDKALEGSYFE 300
Qy 301 MSTYSTNTRMIRILLAFVSSVGRNDTGYTCCSSKHPQSALVTILEKGFINATSSQEEY 360
Db 301 MSTYSTNTRMIRILLAFVSSVGRNDTGYTCCSSKHPQSALVTILEKGFINATSSQEEY 360
Qy 361 EIDPYEKFCFSVRKAYPRICTWIFPSQASPPCEORGLDGYISIKFCDHKKPKGEIIFY 420
Db 361 EIDPYEKFCFSVRKAYPRICTWIFPSQASPPCEORGLDGYISIKFCDHKKPKGEIIFY 420
Qy 421 AENDDAQFTKMTLIRKKPOVLANASASQSDGYPLPSWTWKCKSDKSPNCTEIP 480
Db 421 AENDDAQFTKMTLIRKKPOVLANASASQSDGYPLPSWTWKCKSDKSPNCTEIP 480
Qy 481 EGVNKKANRKYFGQWSSSTLNSEAGGLLVKCCAYNSMGTSCETIFLNSPGPPFIQ 540
Db 481 EGVNKKANRKYFGQWSSSTLNSEAGGLLVKCCAYNSMGTSCETIFLNSPGPPFIQ 540
Qy 541 DNISFYATIGLCPFIIVLVILCHIKYKQFRIESQLOMIQVTPGLDNEYFYVDFRDY 600
Db 541 DNISFYATIGLCPFIIVLVILCHIKYKQFRIESQLOMIQVTPGLDNEYFYVDFRDY 600
Qy 601 DLKWEFPRENLEFGKVLGSGAFGRVWNATAYIGSTGVSIQVAVKMLKEKADSCKEALM 660
Db 601 DLKWEFPRENLEFGKVLGSGAFGRVWNATAYIGSTGVSIQVAVKMLKEKADSCKEALM 660
Qy 661 SELKQWTHLGHHDNIIVNLGACTLSPVYLIFEYCCYDGLNLYLSKREKPHRTWTEIPK 720
Db 661 SELKQWTHLGHHDNIIVNLGACTLSPVYLIFEYCCYDGLNLYLSKREKPHRTWTEIPK 720
Qy 721 EHNFSYPTFOAHSSNMPGSRVQLHPPLDQSGFNNGSIHSEDEIEYENOKRLAEBEE 780
Db 721 EHNFSYPTFOAHSSNMPGSRVQLHPPLDQSGFNNGSIHSEDEIEYENOKRLAEBEE 780

QY 781 EDNLVLTEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVKVICKDFGLARDIL 840
DB 781 EDNLVLTEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVKVICKDFGLARDIL 840
QY 841 SDSSVVRGNARLPVKWMAPELFLFGIYTIKSDVMSYGILLWEIFSLGNVPYGPVDAN 900
DB 841 SDSSVVRGNARLPVKWMAPELFLFGIYTIKSDVMSYGILLWEIFSLGNVPYGPVDAN 900
QY 901 FYKLTQSGFKMBOPFYATGIVFMQSCWAPDSRKRPPFNLTSLFGCOLAEEACIRT 960
DB 901 FYKLTQSGFKMBOPFYATGIVFMQSCWAPDSRKRPPFNLTSLFGCOLAEEACIRT 960
QY 961 SIHLFKQAAPQORGLRAQSPORQVKIHRERS 992
DB 961 SIHLFKQAAPQORGLRAQSPORQVKIHRERS 992

RESULT 6

US-08-601-891-2
; Sequence 2, Application US/08601891
; Patent No. 5747651
; GENERAL INFORMATION:
; APPLICANT: Lemischka, Ihor R.
; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL
; TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ImClone Systems Incorporated
; STREET: 180 Varick Street
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10014
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/601,891
; FILING DATE: 15-FEB-1996
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/977,451
; FILING DATE: 19-NOV-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/906,397
; FILING DATE: 26-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US PCT/US92/05401
; FILING DATE: 26-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: TW 81102961
; FILING DATE: 15-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US PCT/US92/02750
; FILING DATE: 02-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/813,593
; FILING DATE: 24-DEC-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/793,065
; FILING DATE: 15-NOV-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/728,913
; FILING DATE: 28-JUN-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/679,666
; FILING DATE: 02-APR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Felt, Irving N.
; REGISTRATION NUMBER: 28,601

; REFERENCE/DOCKET NUMBER: LEM-3-7P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-645-1405
; TELEFAX: 212-645-2054
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 992 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-601-891-2

Query Match 100.0%; Score 5264; DB 1; Length 992;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 992; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRALAQRSDRRLLLVLSVMILETVTNQDLPVKICVLISHENNGSSACKPSYRMVRS 60
DB 1 MRALAQRSDRRLLLVLSVMILETVTNQDLPVKICVLISHENNGSSACKPSYRMVRS 60
QY 61 PEDLQCTPRQSEGTVEAATVEVAESGSIITLQVLATPGDLSCLWFKHSLGCPHF 120
DB 61 PEDLQCTPRQSEGTVEAATVEVAESGSIITLQVLATPGDLSCLWFKHSLGCPHF 120
QY 121 LQNRGIVSMALNVTTQAGEYLLHIQSERANYTVLFTVNVDRDTQLYVLRPPYFRKXNQ 180
DB 121 LQNRGIVSMALNVTTQAGEYLLHIQSERANYTVLFTVNVDRDTQLYVLRPPYFRKXNQ 180
QY 181 DALLCISGVPEPTVEWVLCSSHRESCKEKGPAVVRKEKVLHELPGTDIRCCARNALGR 240
DB 181 DALLCISGVPEPTVEWVLCSSHRESCKEKGPAVVRKEKVLHELPGTDIRCCARNALGR 240
QY 241 ECTKLTITDLNQAPOSTLPQLFKVGEPLWIRCKAIHVNHGFGLTWELEDKALEESYFE 300
DB 241 ECTKLTITDLNQAPOSTLPQLFKVGEPLWIRCKAIHVNHGFGLTWELEDKALEESYFE 300
QY 301 MSTYSTNRTMIRILLAFVSSVGRNDTGYTCSHSPQSALVTILEKGFINATSQEY 360
DB 301 MSTYSTNRTMIRILLAFVSSVGRNDTGYTCSHSPQSALVTILEKGFINATSQEY 360
QY 361 EIDPYEKECFSVRFKAYPRICTWIFSQASPFCEORGLDGYISIKFCDHKPKGYIFY 420
DB 361 EIDPYEKECFSVRFKAYPRICTWIFSQASPFCEORGLDGYISIKFCDHKPKGYIFY 420
QY 421 AENDDAQFTKMTNIRKKPOVLANASASQASCSDDGYPLPSWTWKKCDKSPNCTEIP 480
DB 421 AENDDAQFTKMTNIRKKPOVLANASASQASCSDDGYPLPSWTWKKCDKSPNCTEIP 480
QY 481 EGVWNNKANRVFGQWSSSTLNMSEAGKLLVKKCAVNSMGTSCTETIFLNSPGPPFIQ 540
DB 481 EGVWNNKANRVFGQWSSSTLNMSEAGKLLVKKCAVNSMGTSCTETIFLNSPGPPFIQ 540
QY 541 DNISFYATIGLCLPFIIVLILCHIKYKQPFYESOLOMIQVTGPDLDNEYVDFRDY 600
DB 541 DNISFYATIGLCLPFIIVLILCHIKYKQPFYESOLOMIQVTGPDLDNEYVDFRDY 600
QY 601 DLKWEFPRENLEFGKVLGSGAFGRVNNATAYGISKTVSIOQAVKMLKEKADSCKEALM 660
DB 601 DLKWEFPRENLEFGKVLGSGAFGRVNNATAYGISKTVSIOQAVKMLKEKADSCKEALM 660
QY 661 SELKMWTHLGHHDNIVNLGACTLSGPVYLIFEYCCYGDLLNLYLSRKRFHRTWTETPK 720
DB 661 SELKMWTHLGHHDNIVNLGACTLSGPVYLIFEYCCYGDLLNLYLSRKRFHRTWTETPK 720
QY 721 EHNFSYPTFOAHNSMPGSEVOLHPPLDOLSGFNGNSIHSEDEIEYENQRLAEEEE 780
DB 721 EHNFSYPTFOAHNSMPGSEVOLHPPLDOLSGFNGNSIHSEDEIEYENQRLAEEEE 780
QY 781 EDNLVLTEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVKVICKDFGLARDIL 840
DB 781 EDNLVLTEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVKVICKDFGLARDIL 840
QY 841 SDSSVVRGNARLPVKWMAPELFLFGIYTIKSDVMSYGILLWEIFSLGNVPYGPVDAN 900

Db 841 SDSSVVRGNARLPVKWMAPESLFEGITKSDVMSYGILLWEIFSLGWNPPGIPVDAN 900
Qy 901 FYKLIQSGFKMEQPEYATGEGYFVWQSCWAFDSRKRPSFPNLTSLGCOLAEAEACIRT 960
Db 901 FYKLIQSGFKMEQPEYATGEGYFVWQSCWAFDSRKRPSFPNLTSLGCOLAEAEACIRT 960
Qy 961 SIHLPKQAAPQOQGLRAQSPQKVIHRS 992
Db 961 SIHLPKQAAPQOQGLRAQSPQKVIHRS 992

RESULT 7

US-09-021-324-2
Sequence 2, Application US/09021324
Patent No. 5912133
GENERAL INFORMATION:
APPLICANT: Lemischka, Ihor R.
TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL
TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: ImClone Systems Incorporated
STREET: 180 Varick Street
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10014
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/021.324
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/977.451
FILING DATE: 1992-11-19
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/906.397
FILING DATE: 26-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US PCT/US92/05401
FILING DATE: 26-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: TW 81102961
FILING DATE: 15-APR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US PCT/US92/02750
FILING DATE: 02-APR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/813.593
FILING DATE: 24-DEC-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/793.065
FILING DATE: 15-NOV-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/728.913
FILING DATE: 28-JUN-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/679.666
FILING DATE: 02-APR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Feit, Irving N.
REGISTRATION NUMBER: 28,601
REFERENCE/DOCKET NUMBER: LEM-3-7P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-645-1405
TELEFAX: 212-645-2054
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:

LENGTH: 992 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-021-324-2
Query Match 100.0%; Score 5264; DB 2; Length 992;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 992; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MRALAQRSDRRLLLVLSVMILETTVTTNQDLPIVKVLISHENNGSSACKPSYRMVGRS 60
Db 1 MRALAQRSDRRLLLVLSVMILETTVTTNQDLPIVKVLISHENNGSSACKPSYRMVGRS 60
Qy 61 PEDLOCTPRROSEGTVEAATVEVAESGITLOVOLATPGDLSCLVFKHSLGCGCPHF 120
Db 61 PEDLOCTPRROSEGTVEAATVEVAESGITLOVOLATPGDLSCLVFKHSLGCGCPHF 120
Qy 121 LQNRGIVSMALNVTETQAGEYLLHIQSERANYTVLFTVNVVDTOLYVLRPFRKXWENQ 180
Db 121 LQNRGIVSMALNVTETQAGEYLLHIQSERANYTVLFTVNVVDTOLYVLRPFRKXWENQ 180
Qy 181 DALLCISEGVPEPTVEWVLCSSHRESCKEEGPAVVRKEEKVLHFLGTDIRCARNALGR 240
Db 181 DALLCISEGVPEPTVEWVLCSSHRESCKEEGPAVVRKEEKVLHFLGTDIRCARNALGR 240
Qy 241 ECTKLFTIDLNAQPOSTLPQLFKVGBEPLWIRCKAIHVNHGFLGTWELDKALEGSYPE 300
Db 241 ECTKLFTIDLNAQPOSTLPQLFKVGBEPLWIRCKAIHVNHGFLGTWELDKALEGSYPE 300
Qy 301 MSTYSTNRTMIRILLAFVSSVGRNDTGYTCTSSSHRPSQSALVTILEKGFINATSSOEY 360
Db 301 MSTYSTNRTMIRILLAFVSSVGRNDTGYTCTSSSHRPSQSALVTILEKGFINATSSOEY 360
Qy 361 EIDPYEKFCSVRFKAYPRIRCTWIFSOASPCQORGLDGYISIKFCDHKNKPGYIYF 420
Db 361 EIDPYEKFCSVRFKAYPRIRCTWIFSOASPCQORGLDGYISIKFCDHKNKPGYIYF 420
Qy 421 AENDDAQTKMFTLNIRKKQVOLANASASQSCSSDGYPLPSWTWKCKSDKSPNCTEIP 480
Db 421 AENDDAQTKMFTLNIRKKQVOLANASASQSCSSDGYPLPSWTWKCKSDKSPNCTEIP 480
Qy 481 EGVNKKANRVFGQWSSSTLNMSSEAGKLLVKCCAYNSMGTSCETIFLNSPGPPFTQ 540
Db 481 EGVNKKANRVFGQWSSSTLNMSSEAGKLLVKCCAYNSMGTSCETIFLNSPGPPFTQ 540
Qy 541 DNISFYATIGLCPFIWVLI VLI CHKYKQFYESQLOMIQVTPGLDNEYFYVDPDRDEY 600
Db 541 DNISFYATIGLCPFIWVLI VLI CHKYKQFYESQLOMIQVTPGLDNEYFYVDPDRDEY 600
Qy 601 DLKWEFPRENLEFGKVLGSGAFGRVMNATAYGISKTGYSIQVAVKMLKEKADSCKEALM 660
Db 601 DLKWEFPRENLEFGKVLGSGAFGRVMNATAYGISKTGYSIQVAVKMLKEKADSCKEALM 660
Qy 661 SELKQWTHLGHNDINVLGACTLSGPVYLIFCYCYGDLLNLRKREKFKHRTWTEIFK 720
Db 661 SELKQWTHLGHNDINVLGACTLSGPVYLIFCYCYGDLLNLRKREKFKHRTWTEIFK 720
Qy 721 EHNFSSTPTFOAHNSNMPGSRREVQLHPPLDQSGFNNSIHSDEIEYENOKRLAEEREE 780
Db 721 EHNFSSTPTFOAHNSNMPGSRREVQLHPPLDQSGFNNSIHSDEIEYENOKRLAEEREE 780
Qy 781 EDLNLVTPEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVKICDFGLARDIL 840
Db 781 EDLNLVTPEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVKICDFGLARDIL 840
Qy 841 SDSSVVRGNARLPVKWMAPESLFEGITKSDVMSYGILLWEIFSLGWNPPGIPVDAN 900
Db 841 SDSSVVRGNARLPVKWMAPESLFEGITKSDVMSYGILLWEIFSLGWNPPGIPVDAN 900
Qy 901 FYKLIQSGFKMEQPEYATGEGYFVWQSCWAFDSRKRPSFPNLTSLGCOLAEAEACIRT 960
Db 901 FYKLIQSGFKMEQPEYATGEGYFVWQSCWAFDSRKRPSFPNLTSLGCOLAEAEACIRT 960

QY 961 SIHLPKQAAPQORGLRAQSPQORQVKKIHRERS 992
DB 961 SIHLPKQAAPQORGLRAQSPQORQVKKIHRERS 992

RESULT 8
PCT-US92-02750-2
; Sequence 2, Application PC/TUS9202750
; GENERAL INFORMATION:
; APPLICANT: LEWISCHKA, IHOR R.
; TITLE OF INVENTION: Totipotent Hematopoietic Stem Cell
; TITLE OF INVENTION: Receptors And Their Ligands
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: IMCLONE SYSTEMS INCORPORATED
; STREET: 180 VARICK STREET
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: US
; ZIP: 10014
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/02750
; FILING DATE: 19920402
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: FEIT, IRVING N.
; REGISTRATION NUMBER: 28,601
; REFERENCE/DOCKET NUMBER: LEM-3-PPPT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-645-1405
; TELEFAX: 212-645-2054
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 992 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US92-02750-2

Query Match 100.0%; Score 5264; DB 5; Length 992;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 992; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRALQSRDRRLLLVLSVMILETVTNQDLPVVKVLSHENNGSSAGKPSYSYRMVGRS 60
DB 1 MRALQSRDRRLLLVLSVMILETVTNQDLPVVKVLSHENNGSSAGKPSYSYRMVGRS 60

QY 61 PEDLQCTPRRQSEGTVEAATVEAEGSITLQVQLATPGDLSCLWFKHSLGCGPHD 120
DB 61 PEDLQCTPRRQSEGTVEAATVEAEGSITLQVQLATPGDLSCLWFKHSLGCGPHD 120

QY 121 LQNRGIVSMAILNVETQAGEYLLHIQSERANYTLFTVNVVDTQYLVLRPFYRMENQ 180
DB 121 LQNRGIVSMAILNVETQAGEYLLHIQSERANYTLFTVNVVDTQYLVLRPFYRMENQ 180

QY 181 DALLCISEGVPEPTVEWVLCSSHRSCKEKGPAVVRKEKVLHFGTDIRCARNALGR 240
DB 181 DALLCISEGVPEPTVEWVLCSSHRSCKEKGPAVVRKEKVLHFGTDIRCARNALGR 240

QY 241 ECKTLFTIDLNQAPSTLPQLFLKVGEPWIRCKAIHVNHHGFLTWELDKALEBGSYFE 300
DB 241 ECKTLFTIDLNQAPSTLPQLFLKVGEPWIRCKAIHVNHHGFLTWELDKALEBGSYFE 300

QY 301 MSTYSTNRTMIRILLAFVSSVGRNDTGYTCTSSSKHPSQSALVTILEKGFINATSSQEBY 360
DB 301 MSTYSTNRTMIRILLAFVSSVGRNDTGYTCTSSSKHPSQSALVTILEKGFINATSSQEBY 360

QY 361 EIDPYEKFCFSVRFKAYPRIRCTWIFSQASFPCEORGLEDSYISKFCDHKKKPGYIFY 420
DB 361 EIDPYEKFCFSVRFKAYPRIRCTWIFSQASFPCEORGLEDSYISKFCDHKKKPGYIFY 420

QY 421 AENDDAOFTKMTLNIRKKPOVLANASASQASCSGDPPLPSWTWKKCSDKSPNCTEETIP 480
DB 421 AENDDAOFTKMTLNIRKKPOVLANASASQASCSGDPPLPSWTWKKCSDKSPNCTEETIP 480

QY 481 EGVWNNKANRVFGQWSSSTLNMSEAGKGLLVKCCAVNSMGTSCETIFLNSPGPPFFIQ 540
DB 481 EGVWNNKANRVFGQWSSSTLNMSEAGKGLLVKCCAVNSMGTSCETIFLNSPGPPFFIQ 540

QY 541 DNISFYATIGLCPPIVVLIVLICHYKQPFYESQLOMIQVTGPDLDNEYFYVDFRDY 600
DB 541 DNISFYATIGLCPPIVVLIVLICHYKQPFYESQLOMIQVTGPDLDNEYFYVDFRDY 600

QY 601 DLKWEPPRENLEFGKVLGSGAPGRVMNATAYGISTGTGYSIOAVXMKLKEKADSCKEALM 660
DB 601 DLKWEPPRENLEFGKVLGSGAPGRVMNATAYGISTGTGYSIOAVXMKLKEKADSCKEALM 660

QY 661 SELKQWTHLGHNDINVLIGACTLSGPVYLIFEYCYGDLNLYLSKREKFKHRTWTEIFK 720
DB 661 SELKQWTHLGHNDINVLIGACTLSGPVYLIFEYCYGDLNLYLSKREKFKHRTWTEIFK 720

QY 721 EHNFSYPTFOAHSNSMPSGREGVQLHPPLDQLSGFNNGSIHSEDEIEYENOKRLAEBEE 780
DB 721 EHNFSYPTFOAHSNSMPSGREGVQLHPPLDQLSGFNNGSIHSEDEIEYENOKRLAEBEE 780

QY 781 EDNLVLTFTEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLTGHGKVKICDFGLARDIL 840
DB 781 EDNLVLTFTEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLTGHGKVKICDFGLARDIL 840

QY 841 SDSSYVVRGNARLPVKWMAPELSLPEGIYTIKSDVMSYGILLWEIFSLGVNYPYGPVDAN 900
DB 841 SDSSYVVRGNARLPVKWMAPELSLPEGIYTIKSDVMSYGILLWEIFSLGVNYPYGPVDAN 900

QY 901 FYKLQSGFKMEQPPYATGTYFVMSQSWAFDRSRKPSFPNLTSLFGCOLAEBEACIRT 960
DB 901 FYKLQSGFKMEQPPYATGTYFVMSQSWAFDRSRKPSFPNLTSLFGCOLAEBEACIRT 960

QY 961 SIHLPKQAAPQORGLRAQSPQORQVKKIHRERS 992
DB 961 SIHLPKQAAPQORGLRAQSPQORQVKKIHRERS 992

RESULT 9
PCT-US92-05401-2
; Sequence 2, Application PC/TUS9205401
; GENERAL INFORMATION:
; APPLICANT: Lemischka, Ihor R.
; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL
; TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: IMCLONE SYSTEMS INCORPORATED
; STREET: 180 VARICK STREET
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: U.S.A.
; ZIP: 10014
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/05401
; FILING DATE: 19920626
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Feit, Irving N.
; REGISTRATION NUMBER: 28,601
; REFERENCE/DOCKET NUMBER: LEM-3-PPPT

TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-645-1405
TELEFAX: 212-645-2054
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 992 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US92-05401-2

Query Match 100.0%; Score 5264; DB 5; Length 992;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 992; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRALQSRDRRLLLVLSMILETVNODLPVVKVLISHENNGSSAGKPPSYRMVGRS 60
DB 1 MRALQSRDRRLLLVLSMILETVNODLPVVKVLISHENNGSSAGKPPSYRMVGRS 60

QY 61 PEDLOCTPRQSEGTVEAATVEVAESGSIITQVQLATPGDLSCLVWFKHSSLGCCQPHFD 120
DB 61 PEDLOCTPRQSEGTVEAATVEVAESGSIITQVQLATPGDLSCLVWFKHSSLGCCQPHFD 120

QY 121 LQNRGIVSMALNVETQAGEYLLHIQSERANTVLTNNVDTQYVLRPFRKXWENQ 180
DB 121 LQNRGIVSMALNVETQAGEYLLHIQSERANTVLTNNVDTQYVLRPFRKXWENQ 180

QY 181 DALLCISEGVPEPTVEWVLCSSHRESCKEAGPAVVRKEEKLHELFGTDIRCCARNALGR 240
DB 181 DALLCISEGVPEPTVEWVLCSSHRESCKEAGPAVVRKEEKLHELFGTDIRCCARNALGR 240

QY 241 ECTKLTIDLNQAPQSTLPQLFKVGEPLWIRCKAIHVNHGFGLTWELEDKALEEGSYFE 300
DB 241 ECTKLTIDLNQAPQSTLPQLFKVGEPLWIRCKAIHVNHGFGLTWELEDKALEEGSYFE 300

QY 301 MSTYSTNRTMIRLLAFVSGVGRNDGYTTCSSSKHPSQSALVTILEKGFNATSSOEY 360
DB 301 MSTYSTNRTMIRLLAFVSGVGRNDGYTTCSSSKHPSQSALVTILEKGFNATSSOEY 360

QY 361 EIDPYEKFCFSVFKAYPIRCTWIFSQSFCEQRLGDEGYSISFKCDHKNGKPGYIFY 420
DB 361 EIDPYEKFCFSVFKAYPIRCTWIFSQSFCEQRLGDEGYSISFKCDHKNGKPGYIFY 420

QY 421 AENDDAQFTMTNIRKPKQVLANASASQSCSDGYPLPSMTWKCKSDKSNCTEETIP 480
DB 421 AENDDAQFTMTNIRKPKQVLANASASQSCSDGYPLPSMTWKCKSDKSNCTEETIP 480

QY 481 EGVWNNKANRKGWGVSSSTLNMSEAGKLLVCCAYNSMGTSCTIFLNSGPPPPFIQ 540
DB 481 EGVWNNKANRKGWGVSSSTLNMSEAGKLLVCCAYNSMGTSCTIFLNSGPPPPFIQ 540

QY 541 DNISFYATIGLCLPFTIWLIVLICHYKQFYESOLOMIQVTGPDNDYFYVDRDYEY 600
DB 541 DNISFYATIGLCLPFTIWLIVLICHYKQFYESOLOMIQVTGPDNDYFYVDRDYEY 600

QY 601 DLKWEFPRENLEBFGKVLGSGAFGRVNNATAYGISTGVSIQVAVKMLKEKADSCKEALM 660
DB 601 DLKWEFPRENLEBFGKVLGSGAFGRVNNATAYGISTGVSIQVAVKMLKEKADSCKEALM 660

QY 661 SELKMWTHLGHNDINVLGACTLSGPVYLIPEYCCYGDLLNLYLRKSKFHTWTETIFK 720
DB 661 SELKMWTHLGHNDINVLGACTLSGPVYLIPEYCCYGDLLNLYLRKSKFHTWTETIFK 720

QY 721 ENHFSYPTFOAHSNMPGSRVQLHPPDLQSGFNGNSIHSEDEIYENOKRLAEBEE 780
DB 721 ENHFSYPTFOAHSNMPGSRVQLHPPDLQSGFNGNSIHSEDEIYENOKRLAEBEE 780

QY 781 EDNLVLTFFDLFCFAYQVAKMEFLEFKSCVHRDLAARNVLVTHGKVKVXICDFGLARDIL 840
DB 781 EDNLVLTFFDLFCFAYQVAKMEFLEFKSCVHRDLAARNVLVTHGKVKVXICDFGLARDIL 840

QY 841 SDSSYVVRGNARLPVKWMAPESLFEGITYIKSDVMSYGILLWEIFSLGVNYPGIPVDAN 900
DB 841 SDSSYVVRGNARLPVKWMAPESLFEGITYIKSDVMSYGILLWEIFSLGVNYPGIPVDAN 900

DB 841 SDSSYVVRGNARLPVKWMAPESLFEGITYIKSDVMSYGILLWEIFSLGVNYPGIPVDAN 900
QY 901 FYKLIQSGFKMEQPFYATEGIVFMQSCWAFDSRKRPPSPNLTSLFLGCLAEAEACIRT 960
DB 901 FYKLIQSGFKMEQPFYATEGIVFMQSCWAFDSRKRPPSPNLTSLFLGCLAEAEACIRT 960
QY 961 SIHLPKQAAPQORGGGLRAQSPQOVKIHRS 992
DB 961 SIHLPKQAAPQORGGGLRAQSPQOVKIHRS 992

RESULT 10

PCT-US92-09893-2
Sequence 2, Application PC/TUS9209893
GENERAL INFORMATION:
APPLICANT: Lemischka, Ihor R.
TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL
TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: ImClone Systems Incorporated
STREET: 180 Varick Street
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10014
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/09893
FILING DATE: 19921116
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Feit, Irving N.
REGISTRATION NUMBER: 28,601
REFERENCE/DOCKET NUMBER: LEM-3-7PT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-645-1405
TELEFAX: 212-645-2054
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 992 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US92-09893-2

Query Match 100.0%; Score 5264; DB 5; Length 992;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 992; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRALQSRDRRLLLVLSMILETVNODLPVVKVLISHENNGSSAGKPPSYRMVGRS 60
DB 1 MRALQSRDRRLLLVLSMILETVNODLPVVKVLISHENNGSSAGKPPSYRMVGRS 60

QY 61 PEDLOCTPRQSEGTVEAATVEVAESGSIITQVQLATPGDLSCLVWFKHSSLGCCQPHFD 120
DB 61 PEDLOCTPRQSEGTVEAATVEVAESGSIITQVQLATPGDLSCLVWFKHSSLGCCQPHFD 120

QY 121 LQNRGIVSMALNVETQAGEYLLHIQSERANTVLTNNVDTQYVLRPFRKXWENQ 180
DB 121 LQNRGIVSMALNVETQAGEYLLHIQSERANTVLTNNVDTQYVLRPFRKXWENQ 180

QY 181 DALLCISEGVPEPTVEWVLCSSHRESCKEAGPAVVRKEEKLHELFGTDIRCCARNALGR 240
DB 181 DALLCISEGVPEPTVEWVLCSSHRESCKEAGPAVVRKEEKLHELFGTDIRCCARNALGR 240

QY 241 ECTKLTIDLNQAPQSTLPQLFKVGEPLWIRCKAIHVNHGFGLTWELEDKALEEGSYFE 300
DB 241 ECTKLTIDLNQAPQSTLPQLFKVGEPLWIRCKAIHVNHGFGLTWELEDKALEEGSYFE 300

```
301 MSTYSTNRTMIRILLAFVSSVGRNDTGYTTCSSSKHPSQSALVTILEKGFINATSSQEEY 360
;
; FILING DATE:
;
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Hasak, Janet E.
; REGISTRATION NUMBER: 28,616
; REFERENCE/DOCKET NUMBER: 879
; TELEPHONE: 415/225-1896
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1000 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
;
; US-08-222-299-2
;
; Query Match
; Best Local Similarity 96.9%; Score 5102; DB 1; Length 1000;
; Matches 968; Conservative 6; Mismatches 17; Indels 10; Gaps 2;
;
QY 1 MRALQRSRRRLLLVLSVMIETVNTQDLPVICKVLISHENNGSSAGKPPSSYRMVRGS 60
DB 1 MRALQRSRRRLLLVLSVMIETVNTQDLPVICKVLISHENNGSSAGKPPSSYRMVRGS 60
QY 61 PEDLOCTPRQSEGVTEAATVEVAESGSIITLQVQLATPGDLSCLWFKHSSLCGCPHFD 120
DB 61 PEDLOCTPRQSEGVTEAATVEVAESGSIITLQVQLATPGDLSCLWFKHSSLCGCPHFD 120
QY 121 LQNRGIVSAIINVTETQAGEYLLHIQSERANYTLFTVNRDQTQYVLRPFRKMEQ 180
DB 121 LQNRGIVSAIINVTETQAGEYLLHIQSERANYTLFTVNRDQTQYVLRPFRKMEQ 180
QY 181 DALLCISEGVPTVEWVLCSSHRSCKEEGPAVVRKEEKLHELFGTDIRCCARNALGR 240
DB 181 DALLCISEGVPTVEWVLCSSHRSCKEEGPAVVRKEEKLHELFGTDIRCCARNALGR 240
QY 241 ECTKLTIDLNOAPOSTLPQLFKVGEPLWIRCKAIHVNHGFLTWELDKALEEGSYFE 300
DB 241 ECTKLTIDLNOAPOSTLPQLFKVGEPLWIRCKAIHVNHGFLTWELDKALEEGSYFE 300
QY 301 MSTYSTNRTMIRILLAFVSSVGRNDTGYTTCSSSKHPSQSALVTILEKGFINATSSQEEY 360
DB 301 MSTYSTNRTMIRILLAFVSSVGRNDTGYTTCSSSKHPSQSALVTILEKGFINATSSQEEY 360
QY 361 EIDPYEKCFSVRFKAYPRICTWIFSQASFCEQORGLDGYISISKFDHKNKPGYIFY 420
DB 361 EIDPYEKCFSVRFKAYPRICTWIFSQASFCEQORGLDGYISISKFDHKNKPGYIFY 420
QY 421 AENDDAQFTKMTLIRKPKQVLNANASQASQSSDGYPLPSMTWKKCDKSPNCTEEIP 480
DB 421 AENDDAQFTKMTLIRKPKQVLNANASQASQSSDGYPLPSMTWKKCDKSPNCTEEIP 480
QY 481 EGVNKKANRVFGQWVSSSTLNMSEAGKGLLVKCCAYNSMGTSCTIFLNSGPPFFIQ 540
DB 481 EGVNKKANRVFGQWVSSSTLNMSEAGKGLLVKCCAYNSMGTSCTIFLNSGPPFFIQ 540
QY 541 DNISFYATTIGLCLPFTVILVILCHYKQFYESOLOMIQVTPGLDNEYFYVDFRDY 600
DB 541 DNISFYATTIGLCLPFTVILVILCHYKQFYESOLOMIQVTPGLDNEYFYVDFRDY 600
QY 601 DLKWEFFPRENLEFGKVLGSGAFGRVNNATAYGTSKTVSIQVAVKMLKEKADSCKEALM 660
DB 601 DLKWEFFPRENLEFGKVLGSGAFGRVNNATAYGTSKTVSIQVAVKMLKEKADSCKEALM 660
QY 661 SELKMWTHLGHNDINVLGACTLSGPVYLIPEYCCYGDLLNLYRSKREKFRHTWTEIFK 720
DB 661 SELKMWTHLGHNDINVLGACTLSGPVYLIPEYCCYGDLLNLYRSKREKFRHTWTEIFK 720
;
; RESULT 11
; US-08-222-299-2
; Sequence 2, Application US/08222299
; Patent No. 5635388
; GENERAL INFORMATION:
; APPLICANT: Bennett, Brian D.
; APPLICANT: Broz, Susan D.
; APPLICANT: Matthews, William
; APPLICANT: Zeigler, Francis C.
; TITLE OF INVENTION: AGONIST ANTIBODIES AGAINST THE FLK2/FLT3 RECEPTOR AND USES THEREIN
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSES: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)
```

```
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/222,299
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Hasak, Janet E.
; REGISTRATION NUMBER: 28,616
; REFERENCE/DOCKET NUMBER: 879
; TELEPHONE: 415/225-1896
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1000 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
;
; US-08-222-299-2
;
; Query Match
; Best Local Similarity 96.7%; Score 5102; DB 1; Length 1000;
; Matches 968; Conservative 6; Mismatches 17; Indels 10; Gaps 2;
;
QY 1 MRALQRSRRRLLLVLSVMIETVNTQDLPVICKVLISHENNGSSAGKPPSSYRMVRGS 60
DB 1 MRALQRSRRRLLLVLSVMIETVNTQDLPVICKVLISHENNGSSAGKPPSSYRMVRGS 60
QY 61 PEDLOCTPRQSEGVTEAATVEVAESGSIITLQVQLATPGDLSCLWFKHSSLCGCPHFD 120
DB 61 PEDLOCTPRQSEGVTEAATVEVAESGSIITLQVQLATPGDLSCLWFKHSSLCGCPHFD 120
QY 121 LQNRGIVSAIINVTETQAGEYLLHIQSERANYTLFTVNRDQTQYVLRPFRKMEQ 180
DB 121 LQNRGIVSAIINVTETQAGEYLLHIQSERANYTLFTVNRDQTQYVLRPFRKMEQ 180
QY 181 DALLCISEGVPTVEWVLCSSHRSCKEEGPAVVRKEEKLHELFGTDIRCCARNALGR 240
DB 181 DALLCISEGVPTVEWVLCSSHRSCKEEGPAVVRKEEKLHELFGTDIRCCARNALGR 240
QY 241 ECTKLTIDLNOAPOSTLPQLFKVGEPLWIRCKAIHVNHGFLTWELDKALEEGSYFE 300
DB 241 ECTKLTIDLNOAPOSTLPQLFKVGEPLWIRCKAIHVNHGFLTWELDKALEEGSYFE 300
QY 301 MSTYSTNRTMIRILLAFVSSVGRNDTGYTTCSSSKHPSQSALVTILEKGFINATSSQEEY 360
DB 301 MSTYSTNRTMIRILLAFVSSVGRNDTGYTTCSSSKHPSQSALVTILEKGFINATSSQEEY 360
QY 361 EIDPYEKCFSVRFKAYPRICTWIFSQASFCEQORGLDGYISISKFDHKNKPGYIFY 420
DB 361 EIDPYEKCFSVRFKAYPRICTWIFSQASFCEQORGLDGYISISKFDHKNKPGYIFY 420
QY 421 AENDDAQFTKMTLIRKPKQVLNANASQASQSSDGYPLPSMTWKKCDKSPNCTEEIP 480
DB 421 AENDDAQFTKMTLIRKPKQVLNANASQASQSSDGYPLPSMTWKKCDKSPNCTEEIP 480
QY 481 EGVNKKANRVFGQWVSSSTLNMSEAGKGLLVKCCAYNSMGTSCTIFLNSGPPFFIQ 540
DB 481 EGVNKKANRVFGQWVSSSTLNMSEAGKGLLVKCCAYNSMGTSCTIFLNSGPPFFIQ 540
QY 541 DNISFYATTIGLCLPFTVILVILCHYKQFYESOLOMIQVTPGLDNEYFYVDFRDY 600
DB 541 DNISFYATTIGLCLPFTVILVILCHYKQFYESOLOMIQVTPGLDNEYFYVDFRDY 600
QY 601 DLKWEFFPRENLEFGKVLGSGAFGRVNNATAYGTSKTVSIQVAVKMLKEKADSCKEALM 660
DB 601 DLKWEFFPRENLEFGKVLGSGAFGRVNNATAYGTSKTVSIQVAVKMLKEKADSCKEALM 660
QY 661 SELKMWTHLGHNDINVLGACTLSGPVYLIPEYCCYGDLLNLYRSKREKFRHTWTEIFK 720
DB 661 SELKMWTHLGHNDINVLGACTLSGPVYLIPEYCCYGDLLNLYRSKREKFRHTWTEIFK 720
```


Qy 721 EHNFSYPTFOAHNSMPSGSEVQLHPPLDQSGFNGSIHSEDEIEYENOKRLAEBEE 780
 Db 721 EHNFSYPTFOAHNSMPSGSEVQLHPPLDQSGFNGSIHSEDEIEYENOKRLAEBEE 780
 Qy 781 EDNLVTFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVKICDFGLARDIL 840
 Db 781 EDNLVTFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVKICDFGLARDIL 840
 Qy 841 SDSSVVRGNARLPVKWMAPELSEFEGYTIKSDVMSYGILLWEIFSLGWNYPGIPVDAN 900
 Db 841 SDSSVVRGNARLPVKWMAPELSEFEGYTIKSDVMSYGILLWEIFSLGWNYPGIPVDAN 900
 Qy 901 FYKLIQSGFKMEQPFYATEGIFVMSQWAFDSRKRPSFPNLTSLGCOLAEEAEACIRT 960
 Db 901 FYKLIQSGFKMEQPFYATEGIFVMSQWAFDSRKRPSFPNLTSLGCOLAEEAEAMYNQ 960
 Qy 961 -----SIHLPKQAAPQQRGLRAQSPQOVKIHRS 992
 Db 961 MGNVPEHPSIQNRRLPSREAGS-EPPSPQAQVKIHRS 1000

RESULT 12
 US-08-434-878-2
 ; Sequence 2, Application US/08434878
 ; Patent No. 5997865
 ; GENERAL INFORMATION:
 ; APPLICANT: Bennett, Brian D.
 ; APPLICANT: Broz, Susan D.
 ; APPLICANT: Matthews, William
 ; APPLICANT: Zeigler, Francis C.
 ; TITLE OF INVENTION: AGONIST ANTIBODIES AGAINST THE FLK2/FLT3 RECEPTOR AND USES THEREOF
 ; NUMBER OF SEQUENCES: 4
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Genentech, Inc.
 ; STREET: 460 Point San Bruno Blvd
 ; CITY: South San Francisco
 ; STATE: California
 ; COUNTRY: USA
 ; ZIP: 94080
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patin (Genentech)
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/434,878
 ; FILING DATE:
 ; CLASSIFICATION: 424
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER:
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Hasak, Janet E.
 ; REGISTRATION NUMBER: 28,616
 ; REFERENCE/DOCKET NUMBER: 879
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 415/225-1896
 ; TELEFAX: 415/952-9881
 ; TELEX: 910/371-7168
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1000 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear

Query Match 96.9%; Score 5102; DB 2; Length 1000;
 Best Local Similarity 96.7%; Pred. No. 0;
 Matches 968; Conservative 6; Mismatches 17; Indels 10; Gaps 2;
 Qy 1 MRALAQRSDRRLLLVLSVMILETTVNQDLPVICKVLISHENNGSSACKPSSYVMVRGS 60
 |||||

Db 1 MRALAQRSDRRLLLVLSVMILETTVNQDLPVICKVLISHENNGSSACKPSSYVMVRGS 60
 Qy 61 PEDLOCTPRRQSEGVYEAAATVEVAESGSIITLQVQLATPGDLSCLWVFHSSIGCOPHFD 120
 Db 61 PEDLOCTPRRQSEGVYEAAATVEVAESGSIITLQVQLATPGDLSCLWVFHSSIGCOPHFD 120
 Qy 121 LONRGIVSMATLANTVETQAGEVLLHIQSERANYTVLFTVNVVDDTQIYVLRPRYFRKQENQ 180
 Db 121 LONRGIVSMATLANTVETQAGEVLLHIQSEAAANYTVLFTVNVVDDTQIYVLRPRYFRKQENQ 180
 Qy 181 DALLCISEGVPEPTVWVLCSSHRESCKEPAVVRKEKVLHELFGTDIRCARNALGR 240
 Db 181 DALLCISEGVPEPTVWVLCSSHRESCKEPAVVRKEKVLHELFGTDIRCARNALGR 240
 Qy 241 ECTKLFTIDLNAQPOSTLPQLFKVGEPLWIRCKAIHVNHGFGLTWELEDKALEEGSYFE 300
 Db 241 ESTLFTIDLNAQPOSTLPQLFKVGEPLWIRCKAIHVNHGFGLTWELEDKALEEGSYFE 300
 Qy 301 MSTYSTNRMTIRILLAFVSSVGRNDTGYTTCSSSKHPSQSALVITILEKGFNATSSQBEY 360
 Db 301 MSTYSTNRMTIRILLAFVSSVGRNDTGYTTCSSSKHPSQSALVITILEKGFNATSSQBEY 360
 Qy 361 EIDPYEKFCSVRPKAYPRIRCTWIFSOASPPCEORGLDGYISKFCDHKNKPGYEIFY 420
 Db 361 EIDPYEKFCSVRPKAYPRIRCTWIFSOASPPCEORGLDGYISKFCDHKNKPGYEIFY 420
 Qy 421 AENDDAQFTKMTLINIRKKPQVLANASASQASCSGYPPLPSWTWKKCDKSPNCTEIP 480
 Db 421 AENDDAQFTKMTLINIRKKPQVLANASASQASCSGYPPLPSWTWKKCDKSPNCTEIP 480
 Qy 481 EGVNKKANRKFQGVWSSSTLNMSEAGKLLKCCAYNSMGTSCTETIFLNSPGFPPIQ 540
 Db 481 EGVNKKANRKFQGVWSSSTLNMSEAGKLLKCCAYNSMGTSCTETIFLNSPGFPPIQ 540
 Qy 541 DNISFYATIGLCLPFIIVLVLVLIHKYKQFVRESQLOMTQVPLDNEYPYVDFRVEY 600
 Db 541 DNISFYATIGLCLPFIIVLVLVLIHKYKQFVRESQLOMTQVPLDNEYPYVDFRVEY 600
 Qy 601 DLKWEFPRENLEFGKVLGSGAFGRVMNATAYGISKTVSIOQAVMKLKEKADSCKEKALM 660
 Db 601 DLKWEFPRENLEFGKVLGSGAFGRVMNATAYGISKTVSIOQAVMKLKEKADSCKEKALM 660
 Qy 661 SELKMTLGHHDNI VNLGACTLSGPFVYLIPFYCCYGDLLNLYRSKREKPHRTWTBIFK 720
 Db 661 SELKMTLGHHDNI VNLGACTLSGPFVYLIPFYCCYGDLLNLYRSKREKPHRTWTBIFK 720
 Qy 721 EHNFSYPTFOAHNSMPSGSEVQLHPPLDQSGFNGSIHSEDEIEYENOKRLAEBEE 780
 Db 721 EHNFSYPTFOAHNSMPSGSEVQLHPPLDQSGFNGSIHSEDEIEYENOKRLAEBEE 780
 Qy 781 EDNLVTFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVKICDFGLARDIL 840
 Db 781 EDNLVTFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVKICDFGLARDIL 840
 Qy 841 SDSSVVRGNARLPVKWMAPELSEFEGYTIKSDVMSYGILLWEIFSLGWNYPGIPVDAN 900
 Db 841 SDSSVVRGNARLPVKWMAPELSEFEGYTIKSDVMSYGILLWEIFSLGWNYPGIPVDAN 900
 Qy 901 FYKLIQSGFKMEQPFYATEGIFVMSQWAFDSRKRPSFPNLTSLGCOLAEEAEACIRT 960
 Db 901 FYKLIQSGFKMEQPFYATEGIFVMSQWAFDSRKRPSFPNLTSLGCOLAEEAEAMYNQ 960
 Qy 961 -----SIHLPKQAAPQQRGLRAQSPQOVKIHRS 992
 Db 961 MGNVPEHPSIQNRRLPSREAGS-EPPSPQAQVKIHRS 1000

RESULT 13
 PCT-US95-03718-2
 ; Sequence 2, Application PC/TUS9503718
 ; GENERAL INFORMATION:
 ; APPLICANT: GENENTECH, INC.
 ; TITLE OF INVENTION: AGONIST ANTIBODIES AGAINST THE FLK2/FLT3 RECEPTOR AND USES THEREOF

INFORMATION FOR SEQ ID NO: 4:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 993 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 US-08-222-299-4

Query Match 84.3%; Score 4436.5; DB 1; Length 993;
 Best Local Similarity 84.0%; Pred. No. 0;
 Matches 836; Conservative 59; Mismatches 87; Indels 13; Gaps 4;

QY 1 MRALQSDRRLLLVLSVMIETVTDLPVTKVLI SHENNGSSACKSPSSYRMVRS 60
 DB 1 MPALA-RDGGQLPLLVFVSAMIFGTITNODLPVTKVLI SHENNGSSACKSPSSYRMVRS 59

QY 61 PEDLOCTPRROSEGTVEAATVEAEGSITLQVLA TPGLSLCLWVFKHSLGCOPHD 120
 DB 60 PEDLGALRPQSSGTVEAATVEAEGSITLQVLA TPGLSLCLWVFKHSLGCOPHD 119

QY 121 LQNRGIVSMALNVTETQAGEYLLHIQSERANYTVLFTVNVDRDTQYVLRPFYFRKMNQ 180
 DB 120 LQNRGVSMVILKMTETQAGEYLLFIQSEATNYTILFTVSI RNTLLYLRPFYFRKMNQ 179

QY 181 DALLCISEGVPEPTVEVWVLCSSHRESCKECPAVVRKEEKVHLFGTDIRCCARNALGR 240
 DB 180 DALVCISESVEPEVWVLCSSHRESCKECPAVVRKEEKVHLFGTDIRCCARNALGR 239

QY 241 ECTKLFTIDLNQAPQSTLPQLFLKVGEPWIRCKAIHVNHGFGLTWELEDKALEEGSYFE 300
 DB 240 ECTRLFTIDLNQPTTLPQLFLKVGEPWIRCKAIHVNHGFGLTWELEDKALEEGSYFE 299

QY 301 MSTYSTNRTWIRILLAPVSSVGRNDTGYTTCSSSKHPSQSALVTILEKGFINATSSOEY 360
 DB 300 MSTYSTNRTWIRILLAPVSSVGRNDTGYTTCSSSKHPSQSALVTILEKGFINATSSOEY 359

QY 361 EIDPEYKFCFVRKAYPRICRTWISQASFPCEQRLGDEYGISKPCDHKNKPGYIFY 420
 DB 360 EIDQVEEFCFVRKAYPRICRTWISQASFPCEQRLGDEYGISKPCDHKNKPGYIFY 419

QY 421 AENDDAQFTKMTLNIRKPKQVLANASASQSCSDGYPLPSWTWKCKSDKSNCTEETP 480
 DB 420 AENDDAQFTKMTLNIRKPKQVLANASASQSCSDGYPLPSWTWKCKSDKSNCTEETP 479

QY 481 EGWNKANKRVFGQWVSSTLNMSEAGKGLVKKCAVNSMGTSCTETFLNSPGPPFFIQ 540
 DB 480 EGWNKANKRVFGQWVSSTLNMSEAGKGLVKKCAVNSMGTSCTETFLNSPGPPFFIQ 539

QY 541 DNISEVATIGLCLPIVULVILI CHYKQKQRYESOLQMIQVTPGLDNDFYVDFRDEY 600
 DB 540 DNISEVATIGLCLPIVULVILI CHYKQKQRYESOLQMIQVTPGLDNDFYVDFRDEY 599

QY 601 DLKWEFPRENLEFGVLGSGAGRVNNTAYGISKTGYSIQVAVNMLKEKADSCKEALM 660
 DB 600 DLKWEFPRENLEFGVLGSGAGRVNNTAYGISKTGYSIQVAVNMLKEKADSCKEALM 659

QY 661 SELKQMTLGHHDNINVLGACTLSGPVYLIPEYCYGDLNLVLSKREKFKHTWTETPK 720
 DB 660 SELKQMTLGHHDNINVLGACTLSGPVYLIPEYCYGDLNLVLSKREKFKHTWTETPK 719

QY 721 EHNFSYTFQAHNSNMPGREGVOLHPPDLQSLGFGNSIHSDEIEIYENOKRLAEDEE 780
 DB 720 EHNFSYTFQAHNSNMPGREGVOLHPPDLQSLGFGNSIHSDEIEIYENOKRLAEDEE 777

QY 781 EDNLVLPEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLTGKVKICDFGLARDIL 840
 DB 780 EDNLVLPEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLTGKVKICDFGLARDIL 837

QY 841 SDSSYVVRGNARLPVKWMAPELSEGIYTIKSDVMSYGLLWEIFSLGWNYPGIPVDAN 900
 DB 840 SDSSYVVRGNARLPVKWMAPELSEGIYTIKSDVMSYGLLWEIFSLGWNYPGIPVDAN 897

QY 901 FYKLIQSGFQMEQPFYATGFIYFMQSCWAFDSRKRPSFNLTSFLGCOLAEAEACIR 960
 DB 901 FYKLIQSGFQMEQPFYATGFIYFMQSCWAFDSRKRPSFNLTSFLGCOLAEAEACIR 956

DB 898 FYKLIQSGFQMEQPFYATGFIYFMQSCWAFDSRKRPSFNLTSFLGCOLAEAEAMYQ- 956
 QY 961 SIHLPKQAP-----QORGGURASQSPQVQK 986
 DB 957 NVDPVSECPHTYQNRPRPSREMDLGLLSPOAQVE 991

RESULT 15

US-08-434-878-4
 ; Sequence 4, Application US/08434878
 ; Patent No. 5997865
 ; GENERAL INFORMATION:
 ; APPLICANT: Bennett, Brian D.
 ; APPLICANT: Broz, Susan D.
 ; APPLICANT: Matthews, William
 ; APPLICANT: Zeigler, Francis C.
 ; TITLE OF INVENTION: AGONIST ANTIBODIES AGAINST THE FLK2/FLT3 RECEPTOR AND USES
 ; NUMBER OF SEQUENCES: 4
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Genentech, Inc.
 ; STREET: 460 Point San Bruno Blvd
 ; CITY: South San Francisco
 ; STATE: California
 ; COUNTRY: USA
 ; ZIP: 94080
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: patin (Genentech)
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/434, 878
 ; FILING DATE:
 ; CLASSIFICATION: 424
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER:
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Haack, Janet E.
 ; REGISTRATION NUMBER: 28,616
 ; REFERENCE/DOCKET NUMBER: 879
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 415/225-1896
 ; TELEFAX: 415/952-9881
 ; TELEX: 910/371-7168
 ; INFORMATION FOR SEQ ID NO: 4:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 993 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 US-08-434-878-4

Query Match 84.3%; Score 4436.5; DB 2; Length 993;
 Best Local Similarity 84.0%; Pred. No. 0;
 Matches 836; Conservative 59; Mismatches 87; Indels 13; Gaps 4;

QY 1 MRALQSDRRLLLVLSVMIETVTDLPVTKVLI SHENNGSSACKSPSSYRMVRS 60
 DB 1 MPALA-RDGGQLPLLVFVSAMIFGTITNODLPVTKVLI SHENNGSSACKSPSSYRMVRS 59

QY 61 PEDLOCTPRROSEGTVEAATVEAEGSITLQVLA TPGLSLCLWVFKHSLGCOPHD 120
 DB 60 PEDLGALRPQSSGTVEAATVEAEGSITLQVLA TPGLSLCLWVFKHSLGCOPHD 119

QY 121 LQNRGIVSMALNVTETQAGEYLLHIQSERANYTVLFTVNVDRDTQYVLRPFYFRKMNQ 180
 DB 120 LQNRGVSMVILKMTETQAGEYLLFIQSEATNYTILFTVSI RNTLLYLRPFYFRKMNQ 179

QY 181 DALLCISEGVPEPTVEVWVLCSSHRESCKECPAVVRKEEKVHLFGTDIRCCARNALGR 240
 DB 180 DALVCISESVEPEVWVLCSSHRESCKECPAVVRKEEKVHLFGTDIRCCARNALGR 239

QY 241 ECTKLFTIDLNQAPQSTLPQLFLKVGEPWIRCKAIHVNHGFGLTWELEDKALEEGSYFE 300

```
Db      240  ECTRLFTIDLNOTPOTTLPOLFLKVGELPLWIRKAVHNGFLGTWELNKALEEGNYFE 299
Qy      301  MSTYSNTRMIRILLAFVSSVGNDRGYTCSSKHPQSALVTILEKGFINATSSOEY 360
Db      300  MSTYSNTRMIRILLAFVSSVANDRGYTCSSKHPQSALVTIIEKGFINATNSEDY 359
Qy      361  EIDPYEKFCSVRFKAYPRIRCTWIFSQASFPCEQRLGLEDGYISKFCOHKNKPGYIFY 420
Db      360  EIDQYEEFCFSVRFKAYPQIRCTWTSRKSPCEQKGLDNGYSISKFCNHKHQPGYIFH 419
Qy      421  AENDDAQFTMFTLNIRKPKQVLANASASQASCSGDPPLPSWTWKCDKSPNCTEETIP 480
Db      420  AENDDAQFTMFTLNIRKPKQVLAERASQASCSGDPPLPSWTWKCDKSPNCTEETIT 479
Qy      481  EGVWNNKANRKGOWSSSTLNMSEAGLLVKCCAYNSMGTSCTIFLNSPGPPFFIQ 540
Db      480  EGVWNNKANRKGOWSSSTLNMSEAIKGLVKCCAYNSLGTSCETILLNSPGPPFFIQ 539
Qy      541  DNISFYATIGLCLPFIWLVILVLIHCHYKKQFRYESQLOMIQVTGPLDNEYFYVDFRDY 600
Db      540  DNISFYATIGVCLLFIVLVLLIHKYKKQFRYESQLOMVQVTGSSDNEYFYVDFREY 599
Qy      601  DLKWEPPRENLEFGKVLGSAFGVKNATAYGISTGVSIQVAVKMLKEKADSCKEALM 660
Db      600  DLKWEPPRENLEFGKVLGSAFGVKNATAYGISTGVSIQVAVKMLKEKADSSEREALM 659
Qy      661  SELKMMTHLGHNDIYNLLGACTLSGPVYLIIEYCCYGDLLNLYRSKREKPHETWTEIF 720
Db      660  SELKMMTQLGSHENIYNLLGACTLSGPYLIIEFYCCYGDLLNLYRSKREKPHETWTEIF 719
Qy      721  EHNFSYPTFOAHNSGMPGSREVQLHPDLDQSGFNHSEDEIEYENQRLAEEEE 780
Db      720  EHNFSYPTFOAHNSGMPGSREVQIHPSDQISGLHNSFHSEDEIEYENQRL--EEE 777
Qy      781  EDNLNLTFFEDLLCFAYQVAKMEFLBFKSCVHRDLAARNVLTGHKVVKICDFGLARDIL 840
Db      778  EDNLNLTFFEDLLCFAYQVAKMEFLBFKSCVHRDLAARNVLTGHKVVKICDFGLARDIM 837
Qy      841  SDSSYVVRGNARLPVKWMAPESLFEGITYTKSDVNSYGILLWEIFSLGVNPPYGPVDAN 900
Db      838  SDSSYVVRGNARLPVKWMAPESLFEGITYTKSDVNSYGILLWEIFSLGVNPPYGPVDAN 897
Qy      901  FYKLIQSGFKMEOPFYATEGIYFVMSQWAFDSRKPSFNLTSFLGCOLAEAEACIRT 960
Db      898  FYKLIQSGFKMEOPFYATEBEIYIIMQSWAFDSRKPSFNLTSFLGCOLADEAEAMYQ- 956
Qy      961  SIHLPKQAAP-----QORGGLRAQSPQOVK 986
Db      957  NVDGPVSECPHTYQNRPPFSREMDLGLLSPQAQVE 991
```

Search completed: August 26, 2003, 07:29:51
Job time : 26.4892 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 27, 2003, 14:24:01 ; Search time 16319.9 Seconds
(without alignments)
8776.084 Million cell updates/sec

Title: US-09-919-408A-3
Perfect score: 3501
Sequence: 1 CGAGGCGCATCCGAGGCT.....AAAAAAAAAAAAAAAAAAAA 3501

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues
Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl.*

1: gb_ba.*

2: gb_hcg.*

3: gb_in.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sv.*

13: gb_un.*

14: gb_vi.*

15: em_ba.*

16: em_fun.*

17: em_hum.*

18: em_in.*

19: em_ma.*

20: em_om.*

21: em_or.*

22: em_ov.*

23: em_pat.*

24: em_ph.*

25: em_pl.*

26: em_ro.*

27: em_sts.*

28: em_un.*

29: em_vi.*

30: em_htg_hum.*

31: em_htg_inv.*

32: em_htg_other.*

33: em_htg_mus.*

34: em_htg_pln.*

35: em_htg_rdt.*

36: em_htg_mam.*

37: em_htg_vrt.*

38: em_sy.*

39: em_hgo_hum.*

40: em_hgo_mus.*

41: em_hgo_other.*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	3501	100.0	3501	6	AR005212	AR005212 Sequence
2	3501	100.0	3501	6	AR071704	AR071704 Sequence
3	3501	100.0	3501	6	I25170	I25170 Sequence 3
4	3501	100.0	3501	6	I40601	I40601 Sequence 3
5	3475	99.3	3475	6	AX695522	AX695522 Sequence 3
6	3475	99.3	3475	6	I44733	I44733 Sequence 3
7	3475	99.3	3475	9	HSU02687	U02687 Human growt
8	3475	99.3	3475	6	I40109	I40109 Sequence 1
9	3393.8	96.9	3422	9	HSPLT3RTK	Z26652 H.sapiens F
10	3083.2	88.1	3120	6	AR262794	AR262794 Sequence 22
11	3083.2	88.1	3120	6	I44518	I44518 Sequence 22
12	2982	85.2	2982	6	AX695523	AX695523 Sequence 1
13	2274.6	65.0	3521	6	I44732	I44732 Sequence 1
14	2274.6	65.0	3521	10	MMFLT3	X59398 Mouse Flt3
15	2227.2	63.6	3453	6	AR005211	AR005211 Sequence
16	2227.2	63.6	3453	6	AR071703	AR071703 Sequence
17	2227.2	63.6	3453	6	I25169	I25169 Sequence 1
18	2227.2	63.6	3453	6	I40600	I40600 Sequence 1
19	2203	62.9	3426	6	AX695519	AX695519 Sequence
20	2203	62.9	3426	10	MUSFLK2	M64689 Mouse flk-2
21	2073	59.2	2979	6	AX695520	AX695520 Sequence
22	1545.8	44.2	2247	6	BD022336	BD022336 Multi-fun
23	624.4	17.8	1894	6	AR149571	AR149571 Sequence
24	558.8	16.0	76721	9	AL591024	AL591024 Human DNA
25	476.8	13.6	480	9	HOM3RTK	L36162 Homo sapien
26	400	11.4	400	11	G11126	G11126 human Sfs S
27	377.4	10.8	3545	5	AF153446	AF153446 Danio rer
28	340.4	9.7	3691	10	BC043054	BC043054 Mus muscu
29	340.4	9.7	3699	10	BC036343	BC036343 Mus muscu
30	339	9.7	4588	5	OMI417832	AJ417832 Oncorhync
31	336	9.6	3679	10	RRCSP1	X61479 Rat mRNA fo
32	330.4	9.4	2894	4	ECA224643	AJ224643 Equus cab
33	330.4	9.4	2973	4	ECA224642	AJ224642 Equus cab
34	330.2	9.4	4577	14	FCSSMONC	K01643 Feline sarc
35	329.8	9.4	3828	4	CATFMS	J03149 Cat (F.dome
36	327.2	9.3	2894	4	ECA224644	AJ224644 Equus cab
37	326.6	9.3	429	10	MUS3RTK	L36163 Mus musculu
38	325.6	9.3	2894	4	ECA224645	AJ224645 Equus cab
39	324.8	9.3	2924	4	AF099030	AF131209 Trichosur
40	324.6	9.3	2937	4	AF099030	AF099030 Canis fam
41	324.2	9.3	3992	6	AX394288	AX394288 Sequence
42	324.2	9.3	3992	6	AX587822	AX587822 Sequence
43	324.2	9.3	3992	9	HSCFMS	X03663 Human mRNA
44	323.2	9.2	2933	4	AF055037	AF055037 Equus cab
45	323.2	9.2	3575	10	MWCFMS	X68932 M.musculus

ALIGNMENTS

RESULT 1
AR005212
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES

AR005212
Sequence 3 from patent US 5747651.
AR005212
AR005212.1 GI:3966091
Unknown.
Unknown.
Unclassified.
1 (bases 1 to 3501)
Lemischka, I.R.
Antibodies against tyrosine kinase receptor flk-1
Patent: US 5747651-A 3 05-MAY-1998;
Location/Qualifiers

3501 bp
DNA
linear
PAT 04-DEC-1998.

Pred. No. is the number of results predicted by chance to have a

Qy	961	ACCTATTCAAACAAACAGAACTATGATACGGATCTCTGTTTGGCTTTTGTATCATCAGTGGCA	1020
Db	961	ACCTATTCAAACAAACAGAACTATGATACGGATCTCTGTTTGGCTTTTGTATCATCAGTGGCA	1020
Qy	1021	AGAAACGACACCGGATACTACCTGCTTCTTCAAAGCATCCAGTCAATCAGCTTTG	1080
Db	1021	AGAAACGACACCGGATACTACCTGCTTCTTCAAAGCATCCAGTCAATCAGCTTTG	1080
Qy	1081	GTTTACCATCGTAGAAAGGATTTATAAAATGCTTACCAATTCAGTGGAAGATTTATGAAAT	1140
Db	1081	GTTTACCATCGTAGAAAGGATTTATAAAATGCTTACCAATTCAGTGGAAGATTTATGAAAT	1140
Qy	1141	GACCAATATGAAGAGTTTGTCTGTCAGGTTTAAAGCCTACCCAATAATCAGATGT	1200
Db	1141	GACCAATATGAAGAGTTTGTCTGTCAGGTTTAAAGCCTACCCAATAATCAGATGT	1200
Qy	1201	ACGTGGACCTTCTCTCGAAAATCATTTCCCTTGTGAGCAAAAGGCTTTCATTAACGGATAC	1260
Db	1201	ACGTGGACCTTCTCTCGAAAATCATTTCCCTTGTGAGCAAAAGGCTTTCATTAACGGATAC	1260
Qy	1261	AGCATATCCAAGTTTGTGCAATCATTAAGACACAGCCAGGAGATATATATTTCCATGCAGAA	1320
Db	1261	AGCATATCCAAGTTTGTGCAATCATTAAGACACAGCCAGGAGATATATATTTCCATGCAGAA	1320
Qy	1321	AATGATGATGCCCAATTTACAAAATTTTTCAGCTGGAATATTAAGAGGAAACCTCAAGTG	1380
Db	1321	AATGATGATGCCCAATTTACAAAATTTTTCAGCTGGAATATTAAGAGGAAACCTCAAGTG	1380
Qy	1381	CTCGCAGAGCATCGGCAAGTCAGGGCTCTGTTTCTCGATGGATACCCATTTACATCT	1440
Db	1381	CTCGCAGAGCATCGGCAAGTCAGGGCTCTGTTTCTCGATGGATACCCATTTACATCT	1440
Qy	1441	TGGACCTGGAAGAAGTGTTTCAGACAAGTCTCCCACTGCACAGAGAGATCAAGAAGCA	1500
Db	1441	TGGACCTGGAAGAAGTGTTTCAGACAAGTCTCCCACTGCACAGAGAGATCAAGAAGCA	1500
Qy	1501	GTCTGGAATAGAAAGCTTAACAGAAAGTTTGGACAGTGGGTGTCGAGCAGTACTCTTA	1560
Db	1501	GTCTGGAATAGAAAGCTTAACAGAAAGTTTGGACAGTGGGTGTCGAGCAGTACTCTTA	1560
Qy	1561	AACATGAGTGAAGCCATAAAGGGTTCTGGTCAAGTGTCTGTGCATACAATTCCTTGGC	1620
Db	1561	AACATGAGTGAAGCCATAAAGGGTTCTGGTCAAGTGTCTGTGCATACAATTCCTTGGC	1620
Qy	1621	ACATCTTGTGAGACGATCCTTTTAAACTCTCCAGGCCCTTCCCTTTTCATCCAAGACAAC	1680
Db	1621	ACATCTTGTGAGACGATCCTTTTAAACTCTCCAGGCCCTTCCCTTTTCATCCAAGACAAC	1680
Qy	1681	ATCTCAATTCATGCAACAATTTGGTGTGTGTCTCTCTTCATTTGTCGTTTAAACCTGCTTA	1740
Db	1681	ATCTCAATTCATGCAACAATTTGGTGTGTGTCTCTCTTCATTTGTCGTTTAAACCTGCTTA	1740
Qy	1741	ATTTGTCAAGTACAAAAGCAATTTAGGTATGAAAGCCAGCTACAGATGTACAGGTG	1800
Db	1741	ATTTGTCAAGTACAAAAGCAATTTAGGTATGAAAGCCAGCTACAGATGTACAGGTG	1800
Qy	1801	ACGGCTCTCATAGATAATCAGTACTTCTACAGTTGATTTTCAGAGAATATGAATATCATCTC	1860
Db	1801	ACGGCTCTCATAGATAATCAGTACTTCTACAGTTGATTTTCAGAGAATATGAATATCATCTC	1860
Qy	1861	AAATGGAGTTTCCAAGAGAAATTTAGATTTGGGAAGGTACTAGGATCAGGTGCTTTT	1920
Db	1861	AAATGGAGTTTCCAAGAGAAATTTAGATTTGGGAAGGTACTAGGATCAGGTGCTTTT	1920
Qy	1921	GGAAAAGTATGAACGCAACAGCTTATGGAATTTAGCAAAACAGGAGTCTCAATCCAGGTT	1980
Db	1921	GGAAAAGTATGAACGCAACAGCTTATGGAATTTAGCAAAACAGGAGTCTCAATCCAGGTT	1980
Qy	1981	GCCGTCAAATGCTGAACGAAAGAGACAGCTCTGAAAGAGAGGCACTCATGTTCAGAA	2040
Db	1981	GCCGTCAAATGCTGAACGAAAGAGACAGCTCTGAAAGAGAGGCACTCATGTTCAGAA	2040
Qy	2041	CTCAAGATGATACCCAGCTGGGAAGCCACGAGAAATTTGTGAACTCTGTGGGGCGTGCTC	2100

Db 301 GAAGTGGATGATCTGCTTCCATCAGCTGCAAGTGTGCTGTCGATGCCCGGGAACATT 360
QY 361 TCCTGTCTCTGGGTCTTTAAGACACAGCTCCCTGGAATGTCAGGCACCAATTTGATTTACAA 420
Db 361 TCCTGTCTCTGGGTCTTTAAGACACAGCTCCCTGGAATGTCAGGCACCAATTTGATTTACAA 420
QY 421 AACAGAGGAGTTGTTTCCATGTGTCATTTTGAATAAGACAGAAACCCCAAGCTGGAGATAC 480
Db 421 AACAGAGGAGTTGTTTCCATGTGTCATTTTGAATAAGACAGAAACCCCAAGCTGGAGATAC 480
QY 481 CTACTTTTTTATTCAGAGTGAAGCTACCAATTAACAATAATGTTTACAGTGAGTATAAGA 540
Db 481 CTACTTTTTTATTCAGAGTGAAGCTACCAATTAACAATAATGTTTACAGTGAGTATAAGA 540
QY 541 AATACCTGCTTTACATTAAGAGACCTTACTTTAGAAAAATGAAAAACAGAGCGCC 600
Db 541 AATACCTGCTTTACATTAAGAGACCTTACTTTAGAAAAATGAAAAACAGAGCGCC 600
QY 601 CTGGTCTGCATATCTGAGAGCTTCAGAGCCGATCGTGGAAATGGGTGCTTTGCCATTC 660
Db 601 CTGGTCTGCATATCTGAGAGCTTCAGAGCCGATCGTGGAAATGGGTGCTTTGCCATTC 660
QY 661 CAGGGGGAAGCTGTAAAGAGAAAGTCCAGCTGTTGTTTAAAAAGGAGAAAAAGTGCTT 720
Db 661 CAGGGGGAAGCTGTAAAGAGAAAGTCCAGCTGTTGTTTAAAAAGGAGAAAAAGTGCTT 720
QY 721 CATGAATTAATTTGGGACGACATAAGTCTGTGCCAGAAATCAACTGGCGAGGGAATGC 780
Db 721 CATGAATTAATTTGGGACGACATAAGTCTGTGCCAGAAATCAACTGGCGAGGGAATGC 780
QY 781 ACCAGGCTGTTACAAATAGATCTAAATCAAACTCTCAGACACACATGTCACCAATTAATT 840
Db 781 ACCAGGCTGTTACAAATAGATCTAAATCAAACTCTCAGACACACATGTCACCAATTAATT 840
QY 841 CTTAAAGTAGGGGAACCCCTTATGGATAAGTGTCAAGCTGTTTCATGTGAACCATGATTC 900
Db 841 CTTAAAGTAGGGGAACCCCTTATGGATAAGTGTCAAGCTGTTTCATGTGAACCATGATTC 900
QY 901 GGGCTCACCTGGGAATAGAAAAACAAAGCACTCGAGGAGGCAACTACTTTGAGATGAT 960
Db 901 GGGCTCACCTGGGAATAGAAAAACAAAGCACTCGAGGAGGCAACTACTTTGAGATGAT 960
QY 961 ACCTATTCAACAAACAGAACTATGATACGGATCTGTTGCTTTGTTATCATCAGTGGA 1020
Db 961 ACCTATTCAACAAACAGAACTATGATACGGATCTGTTGCTTTGTTATCATCAGTGGA 1020
QY 1021 AGAAACGACACCGGATFACTACTTCTCTTCAAGCATCCCAAGTCAATCAGCTTGG 1080
Db 1021 AGAAACGACACCGGATFACTACTTCTCTTCAAGCATCCCAAGTCAATCAGCTTGG 1080
QY 1081 GTTACCATCGTAGGAAAGGATTTATAATGCTTACCAATTTCAAGTGAAGATTAAGAAAT 1140
Db 1081 GTTACCATCGTAGGAAAGGATTTATAATGCTTACCAATTTCAAGTGAAGATTAAGAAAT 1140
QY 1141 GACCAATATGAAGAGTTTGTCTGTCTGTCAGTTTAAAGCTTACCCAAATCAGATGT 1200
Db 1141 GACCAATATGAAGAGTTTGTCTGTCTGTCAGTTTAAAGCTTACCCAAATCAGATGT 1200
QY 1201 AGTGGACCTCTCTCGAAAAATCATTTCTGTGAGCAAAAGGCTCTTCATTAACGGATAC 1260
Db 1201 AGTGGACCTCTCTCGAAAAATCATTTCTGTGAGCAAAAGGCTCTTCATTAACGGATAC 1260
QY 1261 AGCATATCCAAGTTTGTCAATCATAGCACCCAGCCAGGAGATATATATTTCCATGCAGAA 1320
Db 1261 AGCATATCCAAGTTTGTCAATCATAGCACCCAGCCAGGAGATATATATTTCCATGCAGAA 1320
QY 1321 AATGATATGCCCAATTTACAAAATGTTTCAACGCTGAATATAGAAAGGAAACCTCAAGTG 1380
Db 1321 AATGATATGCCCAATTTACAAAATGTTTCAACGCTGAATATAGAAAGGAAACCTCAAGTG 1380
QY 1381 CTCGAGAGCATCCGCAAGTCCGGCTCTGTTCTCGGATGGATACCATTAACCATCT 1440
Db 1381 CTCGAGAGCATCCGCAAGTCCGGCTCTGTTCTCGGATGGATACCATTAACCATCT 1440

QY 1441 TGGACCTGGAAGAGTGTTCAGACAAGTCTCCAACTGCACAGAAGATCACAGAGCA 1500
Db 1441 TGGACCTGGAAGAGTGTTCAGACAAGTCTCCAACTGCACAGAAGATCACAGAGCA 1500
QY 1501 GTCTGGAATAGAAAGGCTAAACAGAAAAAGTGTTCAGACAGTGGGTGTCGAGCAGTACTCTA 1560
Db 1501 GTCTGGAATAGAAAGGCTAAACAGAAAAAGTGTTCAGACAGTGGGTGTCGAGCAGTACTCTA 1560
QY 1561 AACATGAGTGAAGCCATAAAAAGGTTCTCTGTCCTGTCCTTCAATTCGTTTAAACCTGCTTA 1620
Db 1561 AACATGAGTGAAGCCATAAAAAGGTTCTCTGTCCTGTCCTTCAATTCGTTTAAACCTGCTTA 1620
QY 1621 ACATCTGTGAGACGATCCCTTTTAAACTCTCCAGGCCCTTCCCTTTTATCAAGACAAC 1680
Db 1621 ACATCTGTGAGACGATCCCTTTTAAACTCTCCAGGCCCTTCCCTTTTATCAAGACAAC 1680
QY 1681 ATCTCATTTCTATGCAACAATTTGGTGTGTTGTCCTCTCTTCAATTCGTTTAAACCTGCTTA 1740
Db 1681 ATCTCATTTCTATGCAACAATTTGGTGTGTTGTCCTCTCTTCAATTCGTTTAAACCTGCTTA 1740
QY 1741 ATTTGTCAAAAGTACAAAAAGCAATTTAGGTATGAAAGCCAGCTACAGATGTCAGGTTG 1800
Db 1741 ATTTGTCAAAAGTACAAAAAGCAATTTAGGTATGAAAGCCAGCTACAGATGTCAGGTTG 1800
QY 1801 ACCGCTCTCAGATAATGAGTACTTCTACGTTGATTTAGAGAAATATGAATATGATCTC 1860
Db 1801 ACCGCTCTCAGATAATGAGTACTTCTACGTTGATTTAGAGAAATATGAATATGATCTC 1860
QY 1861 AAATGGGAGTTTCCAAAGAGAAATTTAGAGTTTGGAGGTTACTAGGATCAGGTGCTTTT 1920
Db 1861 AAATGGGAGTTTCCAAAGAGAAATTTAGAGTTTGGAGGTTACTAGGATCAGGTGCTTTT 1920
QY 1921 GGAAAGTGTAGCAACGCAACAGCTTTATGAAATTTAGCAAAACAGGAGTCTCAATCCAGTT 1980
Db 1921 GGAAAGTGTAGCAACGCAACAGCTTTATGAAATTTAGCAAAACAGGAGTCTCAATCCAGTT 1980
QY 1981 GCCGTCAAAATGCTGAAAGAAAAAGCAGACAGCTCTGAAAGAGAGGCACTCATGTGAGAA 2040
Db 1981 GCCGTCAAAATGCTGAAAGAAAAAGCAGACAGCTCTGAAAGAGAGGCACTCATGTGAGAA 2040
QY 2041 CTCAGATGATGACCCAGCTGGAGCCAGAGAAATTTGTGAACTGCTGGGGGGCTGC 2100
Db 2041 CTCAGATGATGACCCAGCTGGAGCCAGAGAAATTTGTGAACTGCTGGGGGGCTGC 2100
QY 2101 ACACGTGTCAGGACCAATTTACTTTGATTTTGAATACTGTTGCTATGTTGATCTTCTCAAC 2160
Db 2101 ACACGTGTCAGGACCAATTTACTTTGATTTTGAATACTGTTGCTATGTTGATCTTCTCAAC 2160
QY 2161 TATCTAAGAGTAAAGAGAAAAATTTCAAGGACTTGGACAGAGATTTTCAAGGAACAC 2220
Db 2161 TATCTAAGAGTAAAGAGAAAAATTTCAAGGACTTGGACAGAGATTTTCAAGGAACAC 2220
QY 2221 AATTTTCAGTTTTCACCCACTTCCAAATCAGATCAATCCAAATCCAGCATGCTGTTCAAGA 2280
Db 2221 AATTTTCAGTTTTCACCCACTTCCAAATCAGATCAATCCAAATCCAGCATGCTGTTCAAGA 2280
QY 2281 GAAAGTTTCAAGTACCCCGACTCGGATCAAAATCTCAGGCTTTCATGGGAATTTCAATTTAC 2340
Db 2281 GAAAGTTTCAAGTACCCCGACTCGGATCAAAATCTCAGGCTTTCATGGGAATTTCAATTTAC 2340
QY 2341 TCTGAAGATGAATTTGAATATGAAACCAAAAAAGGCTGGAGAAAGAGAGGAGCTTGAAT 2400
Db 2341 TCTGAAGATGAATTTGAATATGAAACCAAAAAAGGCTGGAGAAAGAGAGGAGCTTGAAT 2400
QY 2401 GTGCTTACATTTGAAGATCTTCTTTGCTTTGATATCAAGTTGCCAAAGGAATGGAATTT 2460
Db 2401 GTGCTTACATTTGAAGATCTTCTTTGCTTTGATATCAAGTTGCCAAAGGAATGGAATTT 2460
QY 2461 CTGGAAATTAAGTCTGCTGTTTCAAGAGACCTGGCCGCCAGGAAAGTCTGTCACCCAC 2520
Db 2461 CTGGAAATTAAGTCTGCTGTTTCAAGAGACCTGGCCGCCAGGAAAGTCTGTCACCCAC 2520

Qy	2521	GGGAAAGTGGTGAAGATATGTGACTTTGGAGTTGGCTCGAGATATCATGAGTGATTCCAAC	2580
Db	2521	GGGAAAGTGGTGAAGATATGTGACTTTGGATTTGGCTCGAGATATCATGAGTGATTCCAAC	2580
Qy	2581	TATGTTGTCAGGGGCAATGCCGCTGCTGCTGTAATAATGGATGGGCCCCGGAAGCGCTGTTT	2640
Db	2581	TATGTTGTCAGGGGCAATGCCGCTGCTGCTGTAATAATGGATGGGCCCCGGAAGCGCTGTTT	2640
Qy	2641	GAAGGCATCTACACCAATTAAGAGTGATGTCTGGTCATATGGAAATATTACTGTGGGAAATC	2700
Db	2641	GAAGGCATCTACACCAATTAAGAGTGATGTCTGGTCATATGGAAATATTACTGTGGGAAATC	2700
Qy	2701	TTCTCACTGGTGGAATCCTTACCTCGGCATTCGGGTGAGTCTAACTCTACAAACTG	2760
Db	2701	TTCTCACTGGTGGAATCCTTACCTCGGCATTCGGGTGAGTCTAACTCTACAAACTG	2760
Qy	2761	ATTCAAAATGGATTTAAAAATGGATCAGCCATTTTATGCTACAGAAGAAATATACATTATA	2820
Db	2761	ATTCAAAATGGATTTAAAAATGGATCAGCCATTTTATGCTACAGAAGAAATATACATTATA	2820
Qy	2821	ATGCAATCCTGCTGGGCTTTTGAATCAAGGAAACGGGCATCCTTCCCTAAATTTGACTTCG	2880
Db	2821	ATGCAATCCTGCTGGGCTTTTGAATCAAGGAAACGGGCATCCTTCCCTAAATTTGACTTCG	2880
Qy	2881	TTTTTTAGATGTCAGCTGGCAGATGTCAGAGAAGCGATGTATCAGATGTGGATGGCCGT	2940
Db	2881	TTTTTTAGATGTCAGCTGGCAGATGTCAGAGAAGCGATGTATCAGATGTGGATGGCCGT	2940
Qy	2941	GTTTTCGGAATGCTCCTCACACCTACAAAACAGGGCGACCTTTCAGCAGAGAGATGGAATTG	3000
Db	2941	GTTTTCGGAATGCTCCTCACACCTACAAAACAGGGCGACCTTTCAGCAGAGAGATGGAATTG	3000
Qy	3001	GGGCTACTCTCCGCGAGGCTCAGGTGGAAGATTCTGTAGAGGAACAATTTAGTTTTAAGG	3060
Db	3001	GGGCTACTCTCCGCGAGGCTCAGGTGGAAGATTCTGTAGAGGAACAATTTAGTTTTAAGG	3060
Qy	3061	ACTTTCATCCCTCACCTATCCCTAACAGGCTGTAGATTACCAAACCAAGATTAATTTTCAT	3120
Db	3061	ACTTTCATCCCTCACCTATCCCTAACAGGCTGTAGATTACCAAACCAAGATTAATTTTCAT	3120
Qy	3121	CACATAAAGAAAATCTATTATCAACTGCTGCTTCCACGAGCTTTTCTCTAGAAGCCGCTCT	3180
Db	3121	CACATAAAGAAAATCTATTATCAACTGCTGCTTCCACGAGCTTTTCTCTAGAAGCCGCTCT	3180
Qy	3181	CGGTTTACTCTGTTTTCAAAAGGACTTTTGTGTAATAATCAATCATCTGTGCAGAAGCGAG	3240
Db	3181	CGGTTTACTCTGTTTTCAAAAGGACTTTTGTGTAATAATCAATCATCTGTGCAGAAGCGAG	3240
Qy	3241	GAGGAGCTGATAATGAACCTTTATTGGAGCATTTGATCTGCATCCAAAGCGCTTCTCAGGCCG	3300
Db	3241	GAGGAGCTGATAATGAACCTTTATTGGAGCATTTGATCTGCATCCAAAGCGCTTCTCAGGCCG	3300
Qy	3301	GCTTGAGTGAATTGTGTAATGTAATGTTTAAAGTCTATGTTTAAATAATATGTATAA	3360
Db	3301	GCTTGAGTGAATTGTGTAATGTAATGTTTAAAGTCTATGTTTAAATAATATGTATAA	3360
Qy	3361	ATTTTGTCTAAGGAGAAGCTAATATGATTTTTTAAAGTCTATGTTTAAATAATATGTATAA	3420
Db	3361	ATTTTGTCTAAGGAGAAGCTAATATGATTTTTTAAAGTCTATGTTTAAATAATATGTATAA	3420
Qy	3421	TTTTTTCAGCTATTTTAAAGTGAATATTTTTTAAAGTCTATGTTTAAATAATATGTATAA	3480
Db	3421	TTTTTTCAGCTATTTTAAAGTGAATATTTTTTAAAGTCTATGTTTAAATAATATGTATAA	3480
Qy	3481	AAAAAAAAAAAAAAAAAAAAA 3501	
Db	3481	AAAAAAAAAAAAAAAAAAAAA 3501	

RESULT 3			
I25170			
LOCUS	I25170	3501 bp	DNA
DEFINITION	Sequence 3 from patent US 5548085.		linear
			PAT 07-OCT-1996

ACCESSION	125170	GI:1605040
VERSION	125170.1	
KEYWORDS		
SOURCE	Unknown.	
ORGANISM	Unknown.	
REFERENCE	Unclassified.	
AUTHORS	1. (bases 1 to 3501)	
TITLE	Lemischka, I.R.	
JOURNAL	Tyrosine kinase receptor human flk-2-specific antibodies	
FEATURES	Patent: US 5548065-A 3 20-AUG-1996;	
source	Location/Qualifiers	
	1..3501	
BASE COUNT	1068 a 709 c 784 g 940 t	
ORIGIN	/organism="unknown"	
Query Match	100.0%; Score 3501; DB 6; Length 3501;	
Best Local Similarity	100.0%; Pred. No. 0;	
Matches 3501; Conservative 0; Mismatches 0; Indels 0; Gaps 0		
Qy	1 CGAGCGCGCATCCGAGGGCTGGGCGCGCGCCCTGGGGACCCCGGGCTCCGGAGGCCCATG 60	
Db	1 CGAGCGCGCATCCGAGGGCTGGGCGCGCGCCCTGGGGACCCCGGGCTCCGGAGGCCCATG 60	
Qy	61 CCGCGCTTGCGCGCGCAGCGGGGCACCGTGCCGTGCTCGTGTGTTTTCTGCAATGATA 120	
Db	61 CCGCGCTTGCGCGCGCAGCGGGGCACCGTGCCGTGCTCGTGTGTTTTCTGCAATGATA 120	
Qy	121 TTTGGGACTATTACAATCAAGATCTGCCGTGTGATCAAGTGTGTTTTTAATCAATCATTAAG 180	
Db	121 TTTGGGACTATTACAATCAAGATCTGCCGTGTGATCAAGTGTGTTTTTAATCAATCATTAAG 180	
Qy	181 AACAAATGATTATCATCAGTGGGGAAGTCATCATATATCCATGTGTATCAGAAATCCCGGAA 240	
Db	181 AACAAATGATTATCATCAGTGGGGAAGTCATCATATATCCATGTGTATCAGAAATCCCGGAA 240	
Qy	241 GACCTCGGGTGTGCGTTTGAGACCCAGAGCTCAGGGAAGTGTACGAAGCTGCCGCTGTG 300	
Db	241 GACCTCGGGTGTGCGTTTGAGACCCAGAGCTCAGGGAAGTGTACGAAGCTGCCGCTGTG 300	
Qy	301 GAAGTGAATGTATCTGCTTCATCACACTGCAAGTGTGTGTCGATGCCCCAGGGAACATT 360	
Db	301 GAAGTGAATGTATCTGCTTCATCACACTGCAAGTGTGTGTCGATGCCCCAGGGAACATT 360	
Qy	361 TCCTGTCTCTGGGTCCTTTAAGCAGAGCTCCCTGAAATTCGACGCGCACCATTTTGATTACA 420	
Db	361 TCCTGTCTCTGGGTCCTTTAAGCAGAGCTCCCTGAAATTCGACGCGCACCATTTTGATTACA 420	
Qy	421 AACAGAGGAGTGTGTTCCATGTGTCATTTTGAAAATGACAGAAACCCAGCTGGAGAAATAC 480	
Db	421 AACAGAGGAGTGTGTTCCATGTGTCATTTTGAAAATGACAGAAACCCAGCTGGAGAAATAC 480	
Qy	481 CTACTTTTTTATTCAGATGAAGCTACCAATTCACAATAATTTGTTTACAGTGAATATAAGA 540	
Db	481 CTACTTTTTTATTCAGATGAAGCTACCAATTCACAATAATTTGTTTACAGTGAATATAAGA 540	
Qy	541 AATACCCCTGCTTTACATTAAGAGACCTTACTTTAGAAAAATGGAACACACGAGCGCC 600	
Db	541 AATACCCCTGCTTTACATTAAGAGACCTTACTTTAGAAAAATGGAACACACGAGCGCC 600	
Qy	601 CTGCTCTGCATATCTGAGAGCGTTTCCAGAGCCGATCGTGGAAATGGGTGCTTTTGGCATTC 660	
Db	601 CTGCTCTGCATATCTGAGAGCGTTTCCAGAGCCGATCGTGGAAATGGGTGCTTTTGGCATTC 660	
Qy	661 CAGGGGAAAGCTGTAAAGAAAGAAAGTCCAGCTGTGTTTAAAAAGGAGGAAAAAGTGCTT 720	
Db	661 CAGGGGAAAGCTGTAAAGAAAGAAAGTCCAGCTGTGTTTAAAAAGGAGGAAAAAGTGCTT 720	
Qy	721 CATGAATTAATTGGGACGGACATAAGTGTGTGTCAGAAATGAACTGGCGCAGGGAATGC 780	
Db	721 CATGAATTAATTGGGACGGACATAAGTGTGTGTCAGAAATGAACTGGCGCAGGGAATGC 780	
Qy	781 ACCAGGCTGTTTCAATATAGATCTTAATCAAACTCCTTCAGACCAACATTTGGCCCAATTTT 840	

[illegible]

1861	DB	AAATGGGAGT	TTCCAAAGAGAAAATTTAGAGTTTGGGAAGGTACTAGAGATCAGGTGCTTTTT	1921
1921	QY	GGAAAAGT	GTGATGAACGCAACAGACTTATGGAATTAGCAAAAACGAGAGTCTCAATCCAGGTT	1980
1921	DB	GGAAAAGT	GTGATGAACGCAACAGACTTATGGAATTAGCAAAAACGAGAGTCTCAATCCAGGTT	1980
1981	QY	GCCGTCAAAT	GTCTGAAGAAGAAAAGCAGACAGCTCTGAAAAGAGGACCATCATGTTCAGAA	2040
1981	DB	GCCGTCAAAT	GTCTGAAGAAGAAAAGCAGACAGCTCTGAAAAGAGGACCATCATGTTCAGAA	2040
2041	QY	CTCAAGATGAT	GCACCCAGCTGGGAAGCCACGAGAAATATCTGTGAACCTGCTGGGGGGCGTGC	2100
2041	DB	CTCAAGATGAT	GCACCCAGCTGGGAAGCCACGAGAAATATCTGTGAACCTGCTGGGGGGCGTGC	2100
2101	QY	ACACTGT	CAGAGCCAAATTTACTTGAATTTTGAATACCTGTGTATGGTGAATCTTCTCAAC	2160
2101	DB	ACACTGT	CAGAGCCAAATTTACTTGAATTTTGAATACCTGTGTATGGTGAATCTTCTCAAC	2160
2161	QY	TATCTAAGA	AGTAAAGAGAAAAATTTCAAGGACTTTGGAAGAGAGATTTTCAAGGAACAC	2220
2161	DB	TATCTAAGA	AGTAAAGAGAAAAATTTCAAGGACTTTGGAAGAGAGATTTTCAAGGAACAC	2220
2221	QY	AAATTTCA	GTCTTACCCCACTTTCCATCACATCCAAATTCACAGCATGCCTGGTTCAGA	2280
2221	DB	AAATTTCA	GTCTTACCCCACTTTCCATCACATCCAAATTCACAGCATGCCTGGTTCAGA	2280
2281	QY	GAACTTC	CAGATCACACCCGGACTCGGATCAAACTCTCAGGGCTTCATGGGAATTCATTTCA	2340
2281	DB	GAACTTC	CAGATCACACCCGGACTCGGATCAAACTCTCAGGGCTTCATGGGAATTCATTTCA	2340
2341	QY	TCTGAAGAT	GAAATTTGAATATGAAAAACCAAAAAAGGCTGGAAAGAGGAGGACTTTGAAT	2400
2341	DB	TCTGAAGAT	GAAATTTGAATATGAAAAACCAAAAAAGGCTGGAAAGAGGAGGACTTTGAAT	2400
2401	QY	GTGCTTAC	ATTATTGAAAGATCTTCTTGGTGTGCATATCAAGTTGCCCAAGGAATGGAATTT	2460
2401	DB	GTGCTTAC	ATTATTGAAAGATCTTCTTGGTGTGCATATCAAGTTGCCCAAGGAATGGAATTT	2460
2461	QY	CTGGAA	TTTTAAGTCGTGTGTTCACAGAGACCTGGCCGCGCAGGAACGTGTGTGCACCCAC	2520
2461	DB	CTGGAA	TTTTAAGTCGTGTGTTCACAGAGACCTGGCCGCGCAGGAACGTGTGTGCACCCAC	2520
2521	QY	GGGAAAGT	GTGAAGATATGTGACTTTTGGATGGCTCGAGATATCATAGATGATTTCCAAC	2580
2521	DB	GGGAAAGT	GTGAAGATATGTGACTTTTGGATGGCTCGAGATATCATAGATGATTTCCAAC	2580
2581	QY	TATGTTGT	CAGGGGCAATGCCGCTCGCTGTAAATGATGGCCCCCGAAGCCCTGTTT	2640
2581	DB	TATGTTGT	CAGGGGCAATGCCGCTCGCTGTAAATGATGGCCCCCGAAGCCCTGTTT	2640
2641	QY	GAAGGCAT	CTACACCATTAAGAGTGATGTCGTGTATATGGAATATTTACTGTGGGAAATC	2700
2641	DB	GAAGGCAT	CTACACCATTAAGAGTGATGTCGTGTATATGGAATATTTACTGTGGGAAATC	2700
2701	QY	TTCTCACT	TTGGTGAATCCTTACCCTGGCAATCCGGTTGATGCTAACTTCTACAAACTG	2760
2701	DB	TTCTCACT	TTGGTGAATCCTTACCCTGGCAATCCGGTTGATGCTAACTTCTACAAACTG	2760
2761	QY	ATTCAA	AAATGGATTTAAATGGAATCAGCCATTTTATGCTACAGAGAATAATACATTATA	2820
2761	DB	ATTCAA	AAATGGATTTAAATGGAATCAGCCATTTTATGCTACAGAGAATAATACATTATA	2820
2821	QY	ATGCAAT	CTCGCTGGGCTTTTGACTCAAGGAACGGCCATCCTTCCCTAATTTGACTTCG	2880
2821	DB	ATGCAAT	CTCGCTGGGCTTTTGACTCAAGGAACGGCCATCCTTCCCTAATTTGACTTCG	2880
2881	QY	TTTTTAG	ATGTCAGATGCAGAGAAGCGATGTATCAGAAATGTGGATGGCCGT	2940
2881	DB	TTTTTAG	ATGTCAGATGCAGAGAAGCGATGTATCAGAAATGTGGATGGCCGT	2940
2941	QY	GTTTCGA	ATGTCCTCACACTTACAAAACAGCGGACCTTTTCAGCAGAGAGATGGAATTCG	3000
2941	DB	GTTTCGA	ATGTCCTCACACTTACAAAACAGCGGACCTTTTCAGCAGAGAGATGGAATTCG	3000

```
QY 3001 GGGCTACTCTCTCCGAGGCTCAGTCCGAAGTTCGTAGAGGAACAATTTAGTTTAAAG 3060
Db 3001 GGGCTACTCTCTCCGAGGCTCAGTCCGAAGTTCGTAGAGGAACAATTTAGTTTAAAG 3060
QY 3061 ACTTCATCCCTCCACCTATCCCTAACAGGCTGTAGATTACCAAAAACAAGATTAAATTCAT 3120
Db 3061 ACTTCATCCCTCCACCTATCCCTAACAGGCTGTAGATTACCAAAAACAAGATTAAATTCAT 3120
QY 3121 CACTAAAGAAATCTATTATCACTCACTGCTTCCAGAGACTTTCTCTAGAGCCGCT 3180
Db 3121 CACTAAAGAAATCTATTATCACTCACTGCTTCCAGAGACTTTCTCTAGAGCCGCT 3180
QY 3181 GCGTTTACTCTGTTTCAAGAGGACTTTTGTAAATCAATCATCTGTCACAGAGCTTTCTCTAGAGCCGCT 3240
Db 3181 GCGTTTACTCTGTTTCAAGAGGACTTTTGTAAATCAATCATCTGTCACAGAGCTTTCTCTAGAGCCGCT 3240
QY 3241 GAGGAGCTGATTAATGAACCTTTATTGGAGCAATGATCTGCATCCAGGCTTTCTCAGGCG 3300
Db 3241 GAGGAGCTGATTAATGAACCTTTATTGGAGCAATGATCTGCATCCAGGCTTTCTCAGGCG 3300
QY 3301 GCTTGAGTGAATGTTGTAACCTGAAGTACAGTATATCTTTGTAATAACATAAAACAAGC 3360
Db 3301 GCTTGAGTGAATGTTGTAACCTGAAGTACAGTATATCTTTGTAATAACATAAAACAAGC 3360
QY 3361 ATTTTGTCTAAGGAGCTATATGATTTTAAAGTCTATCTTTTAAATTAATATGTA 3420
Db 3361 ATTTTGTCTAAGGAGCTATATGATTTTAAAGTCTATCTTTTAAATTAATATGTA 3420
QY 3421 TTTTTCAGCTATTATGATGATATATTTATGGCTGGGAATAAAATTTCTACTACAGAAAA 3480
Db 3421 TTTTTCAGCTATTATGATGATATATTTATGGCTGGGAATAAAATTTCTACTACAGAAAA 3480
QY 3481 AAAAAAAAAAAAAAAAAAAAAA 3501
Db 3481 AAAAAAAAAAAAAAAAAAAAAA 3501

RESULT 4
LOCUS I40601 3501 bp DNA linear PAT 13-MAY-1997
DEFINITION Sequence 3 from patent US 5621090.
ACCESSION I40601
VERSION I40601.1 GI:2082893
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 3501)
AUTHORS Lemischka,I.R.
TITLE Nucleic acids encoding soluble human FLK-2 extracellular domain
JOURNAL Patent: US 5621090-A 3 15-APR-1997;
FEATURES
    Location/Qualifiers
        source
            1..3501
                /organism="unknown"
BASE COUNT 1068 a 709 c 784 g 940 t
ORIGIN
Query Match 100.0%; Score 3501; DB 6; Length 3501;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGAGGGCGCATCCGAGGGCTGGCGCGCCCTGGGGACCCCGGGCTCCGGAGGCCATG 60
Db 1 CGAGGGCGCATCCGAGGGCTGGCGCGCCCTGGGGACCCCGGGCTCCGGAGGCCATG 60
QY 61 CCGGCGTTGGCGCGGACCGGCGCGCCCTGGGGACCCCGGGCTCCGGAGGCCATG 120
Db 61 CCGGCGTTGGCGCGGACCGGCGCGCCCTGGGGACCCCGGGCTCCGGAGGCCATG 120
QY 121 TTTGGGACTATTACAATCAAGATCTGCGCTGTGATCAAGTGTGTTTAAATCAATCATAAG 180
Db 121 TTTGGGACTATTACAATCAAGATCTGCGCTGTGATCAAGTGTGTTTAAATCAATCATAAG 180
```

```
QY 181 AACAAATGATTCATCAGTGGGAAAGTCAATCATATCCATGGTATCAGAATCCCGGAA 240
Db 181 AACAAATGATTCATCAGTGGGAAAGTCAATCATATCCATGGTATCAGAATCCCGGAA 240
QY 241 GACCTCGGCTGCTGGTTTGAGACCCAGAGCTCAGGACAGTGTACGAAGCTGCCCGCTG 300
Db 241 GACCTCGGCTGCTGGTTTGAGACCCAGAGCTCAGGACAGTGTACGAAGCTGCCCGCTG 300
QY 301 GAAGTGAATGATCTGCTTCCATCACTGCAAGTGTGCTGATGCTGCCAGGCAACATT 360
Db 301 GAAGTGAATGATCTGCTTCCATCACTGCAAGTGTGCTGATGCTGCCAGGCAACATT 360
QY 361 TCCTGCTCTCGGCTTTTAAGCAGAGCTCCCTGAAATTCGCGAGCCACATTTTGA 420
Db 361 TCCTGCTCTCGGCTTTTAAGCAGAGCTCCCTGAAATTCGCGAGCCACATTTTGA 420
QY 421 AACAGAGGAGTGTTCCTCATGCTCATTTTGAATAATGACAGAAACCCAGAGCTGGAA 480
Db 421 AACAGAGGAGTGTTCCTCATGCTCATTTTGAATAATGACAGAAACCCAGAGCTGGAA 480
QY 481 CTACTTTTTTATTCAGAGTGAAGCTACCAATTAACAATATTGTTTACAGTGAATAA 540
Db 481 CTACTTTTTTATTCAGAGTGAAGCTACCAATTAACAATATTGTTTACAGTGAATAA 540
QY 541 AATACCTGCTTTTACACATTAAGAAGACCTTACTTTAGAAAAATGGAACCCAGAG 600
Db 541 AATACCTGCTTTTACACATTAAGAAGACCTTACTTTAGAAAAATGGAACCCAGAG 600
QY 601 CTGCTGCTCATATCTGAGAGCGTTCCAGAGCGATCGTGAATGGTCTCTTTCGCA 660
Db 601 CTGCTGCTCATATCTGAGAGCGTTCCAGAGCGATCGTGAATGGTCTCTTTCGCA 660
QY 661 CAGGGGAAAGCTGTAAGAAGAAAGTCCAGCTGTGTTTAAAAAGGAGGAAAGTGT 720
Db 661 CAGGGGAAAGCTGTAAGAAGAAAGTCCAGCTGTGTTTAAAAAGGAGGAAAGTGT 720
QY 721 CATGAATTTATTTGGGACCGACATAAGTGTGTCGCCAGAAATGAATGGGCGGAAATGC 780
Db 721 CATGAATTTATTTGGGACCGACATAAGTGTGTCGCCAGAAATGAATGGGCGGAAATGC 780
QY 781 ACCAGGCTGTTCACATAGATCTAAATCAAACTCCTCAGACCAATTCGCAATTA 840
Db 781 ACCAGGCTGTTCACATAGATCTAAATCAAACTCCTCAGACCAATTCGCAATTA 840
QY 841 CTTAAAGTACGGAAACCTTTATGATAAGTGCAGAAAGCTGTCATGTGAACCATG 900
Db 841 CTTAAAGTACGGAAACCTTTATGATAAGTGCAGAAAGCTGTCATGTGAACCATG 900
QY 901 GGGCTCACCTGGGAAATTAGAAAAACAAAGCACTCGAGGAGGCACTACTTTGAGATG 960
Db 901 GGGCTCACCTGGGAAATTAGAAAAACAAAGCACTCGAGGAGGCACTACTTTGAGATG 960
QY 961 ACCTATTCAACAAACAGAACTATGATACGGATTCTGTTGCTTTGTTATCATCAG 1020
Db 961 ACCTATTCAACAAACAGAACTATGATACGGATTCTGTTGCTTTGTTATCATCAG 1020
QY 1021 AGAAACGACACCGGATACACTGTTGTTCTCTTCAAGCATCCAGTCAATCAGCTTT 1080
Db 1021 AGAAACGACACCGGATACACTGTTGTTCTCTTCAAGCATCCAGTCAATCAGCTTT 1080
QY 1081 GTTACCATTGTTAGAAAGGATTATTAATGCTACCAATTCAGTGAAGATTATGA 1140
Db 1081 GTTACCATTGTTAGAAAGGATTATTAATGCTACCAATTCAGTGAAGATTATGA 1140
QY 1141 GACCAATATGAAGATTTTGTGTTTCTGTGAGCAAAAGGCTTTGATTAACGATAT 1200
Db 1141 GACCAATATGAAGATTTTGTGTTTCTGTGAGCAAAAGGCTTTGATTAACGATAT 1200
QY 1201 ACCTGACCTTCTCTCGAAATCAATTTCTTGTGAGCAAAAGGCTTTGATTAACGATAT 1260
Db 1201 ACCTGACCTTCTCTCGAAATCAATTTCTTGTGAGCAAAAGGCTTTGATTAACGATAT 1260
```

QY 1261 AGCATATCCAAAGTTTGGCAATCATAGCACAGCCAGGAGAAATATATATATTCATGCGAGAA 1320
DB 1261 AGCATATCCAAAGTTTGGCAATCATAGCACAGCCAGGAGAAATATATATATTCATGCGAGAA 1320
QY 1321 AATGATGATGCCCAATTTACCAAAATGTTTCAAGCTGAATATAGAGAAACCTCAAGTG 1380
DB 1321 AATGATGATGCCCAATTTACCAAAATGTTTCAAGCTGAATATAGAGAAACCTCAAGTG 1380
QY 1381 CTCGCAAGAGCATCGGCAAGTCAGGGGCTCTGTTTCTCGGATGGATACCCATACCATCT 1440
DB 1381 CTCGCAAGAGCATCGGCAAGTCAGGGGCTCTGTTTCTCGGATGGATACCCATACCATCT 1440
QY 1441 TGGACCTGGAAGAGTGTTCAGACAAGTCTCCCAAATGCGACAGAGATCAGAGAGGA 1500
DB 1441 TGGACCTGGAAGAGTGTTCAGACAAGTCTCCCAAATGCGACAGAGATCAGAGAGGA 1500
QY 1501 GTCTGGAATAGAAAGGCTAACAGAAAGTGTGGACAGTGGGTGTCGACGACTACTCTA 1560
DB 1501 GTCTGGAATAGAAAGGCTAACAGAAAGTGTGGACAGTGGGTGTCGACGACTACTCTA 1560
QY 1561 AACATGATGAAGCCATAAAAGGGTTCCTGGTCAAGTGTCTGTGCATACAAATTCCTTTGGC 1620
DB 1561 AACATGATGAAGCCATAAAAGGGTTCCTGGTCAAGTGTCTGTGCATACAAATTCCTTTGGC 1620
QY 1621 ACATCTGTGAGACGATCCTTTTAAACTCTCCAGGCCCTTCCCTTTTCATCCCAAGACAAC 1680
DB 1621 ACATCTGTGAGACGATCCTTTTAAACTCTCCAGGCCCTTCCCTTTTCATCCCAAGACAAC 1680
QY 1681 ATCTCAATTCATCAACAATGTGGTGTGTCCTCTTCAATGTCGTTTAAACCTGCTA 1740
DB 1681 ATCTCAATTCATCAACAATGTGGTGTGTCCTCTTCAATGTCGTTTAAACCTGCTA 1740
QY 1741 ATTTGTACAAAGTACAAAAGCAATTTAGGTATGAAGCCAGCTTACAGATGGTACAGGTG 1800
DB 1741 ATTTGTACAAAGTACAAAAGCAATTTAGGTATGAAGCCAGCTTACAGATGGTACAGGTG 1800
QY 1801 ACCGGCTCTCAGATATAGTACTTCTACGTTGATTTTCAGAGAAATGATGATCTC 1860
DB 1801 ACCGGCTCTCAGATATAGTACTTCTACGTTGATTTTCAGAGAAATGATGATCTC 1860
QY 1861 AAATGGAGTTTCCAGAGAAATTTAGCTTTGGGAGGTACTAGGATCAGGTGCTTTT 1920
DB 1861 AAATGGAGTTTCCAGAGAAATTTAGCTTTGGGAGGTACTAGGATCAGGTGCTTTT 1920
QY 1921 GGAAAAGTGTAAACGCAACAGCTTATGGAATTAGCAAAAACAGAGTCTCAATCCAGGTT 1980
DB 1921 GGAAAAGTGTAAACGCAACAGCTTATGGAATTAGCAAAAACAGAGTCTCAATCCAGGTT 1980
QY 1981 GCCGTCAAAATGCTGAAAGAAAAGCAGACAGCTCTGAAAAGAGGCACTCATGTCAAGAA 2040
DB 1981 GCCGTCAAAATGCTGAAAGAAAAGCAGACAGCTCTGAAAAGAGGCACTCATGTCAAGAA 2040
QY 2041 CTCAGATGATGACCGAGCTGGGAGCCAGAGNATATGTAACCTGCTGGGGGCGTGC 2100
DB 2041 CTCAGATGATGACCGAGCTGGGAGCCAGAGNATATGTAACCTGCTGGGGGCGTGC 2100
QY 2101 ACACGTGCAGACCAATTTACTGATTTTGAATGACTGTTGCTATGTTGATCTTCTCAAC 2160
DB 2101 ACACGTGCAGACCAATTTACTGATTTTGAATGACTGTTGCTATGTTGATCTTCTCAAC 2160
QY 2161 TATCTAAGAAAGTAAAGAGAAAAATTTACAGGACTTTGGACAGAGATTTTCAAGGAACAC 2220
DB 2161 TATCTAAGAAAGTAAAGAGAAAAATTTACAGGACTTTGGACAGAGATTTTCAAGGAACAC 2220
QY 2221 AATTTCAAGTTTATCCCACTTTCCCAATCAATCCAAATTCAGCATGCTCGTTCAAGA 2280
DB 2221 AATTTCAAGTTTATCCCACTTTCCCAATCAATCCAAATTCAGCATGCTCGTTCAAGA 2280
QY 2281 GAAGTTTCAGATACACCGGACTCGGATCAATCTCAGGGCTTCATGGGAATTCATTTTCAAC 2340
DB 2281 GAAGTTTCAGATACACCGGACTCGGATCAATCTCAGGGCTTCATGGGAATTCATTTTCAAC 2340
QY 2341, TCTGAAGATGAAATTTGAATATGAAAAACCAAAAAAGGCTGGAAAGAGGAGGACTTGAAT 2400

DB 2341 TCTGAAGATGAAATTTGAATATGAAAAACCAAAAAAGGCTGGAAAGAGGAGGACTTGAAT 2400
QY 2401 GTGCTTACATTTGAAGATCTTCTTTGCTTTGCAATATCAAGTTTCCAAAGGAATGGAAATTT 2460
DB 2401 GTGCTTACATTTGAAGATCTTCTTTGCTTTGCAATATCAAGTTTCCAAAGGAATGGAAATTT 2460
QY 2461 CTGGAATTTAAGTCTGTGTGTTCACAGAGACCTGGCCGCAAGAACTGTCTGTTCACCCAC 2520
DB 2461 CTGGAATTTAAGTCTGTGTGTTCACAGAGACCTGGCCGCAAGAACTGTCTGTTCACCCAC 2520
QY 2521 GGGAAAGTGGTGAAGATATGTGACTTTGGATTTGGCTCGAGATATCATGAGTGAATTCACAAC 2580
DB 2521 GGGAAAGTGGTGAAGATATGTGACTTTGGATTTGGCTCGAGATATCATGAGTGAATTCACAAC 2580
QY 2581 TATGTTGTGAGGGCAATGCCGCTCTGCTCTGTAATAATGAGATGGCCCCGCAAGGCTGTTTT 2640
DB 2581 TATGTTGTGAGGGCAATGCCGCTCTGCTCTGTAATAATGAGATGGCCCCGCAAGGCTGTTTT 2640
QY 2641 GAAGGATCTACACCATTAAGAGTGTCTGTGCATATGGAATATTAATCTGTGGGAAATC 2700
DB 2641 GAAGGATCTACACCATTAAGAGTGTCTGTGCATATGGAATATTAATCTGTGGGAAATC 2700
QY 2701 TTCTCACTTTGGTGTGAATCCTTACCTGGCATTTCCGGTTGATCGTAACCTTCTACAAACTG 2760
DB 2701 TTCTCACTTTGGTGTGAATCCTTACCTGGCATTTCCGGTTGATCGTAACCTTCTACAAACTG 2760
QY 2761 ATTCAAAATGGATTTAAAATGAGTCAAGCCATTTTATGCTACAGAAAGAAATATACATTATA 2820
DB 2761 ATTCAAAATGGATTTAAAATGAGTCAAGCCATTTTATGCTACAGAAAGAAATATACATTATA 2820
QY 2821 ATGCAATCTCTGCTGGGCTTTTGAATCAAGAAAACGGCCATCTCTTCCCTAAATTTGACTTCG 2880
DB 2821 ATGCAATCTCTGCTGGGCTTTTGAATCAAGAAAACGGCCATCTCTTCCCTAAATTTGACTTCG 2880
QY 2881 TTTTATGAGTGTGAGTGGCAGATGCGAGAGAGCGATGTATCAGATGTGGATGGGCGGT 2940
DB 2881 TTTTATGAGTGTGAGTGGCAGATGCGAGAGAGCGATGTATCAGATGTGGATGGGCGGT 2940
QY 2941 GTTTCGGAATGTCTCTCACACCTTACCAAAAACAGCGGACCTTTTACAGCAGAGATGGATTTG 3000
DB 2941 GTTTCGGAATGTCTCTCACACCTTACCAAAAACAGCGGACCTTTTACAGCAGAGATGGATTTG 3000
QY 3001 GGCGTACTCTCTCCGAGGCTCAGGTGGAAGATTCGTAGAGGAACAATTTAGTTTAAAG 3060
DB 3001 GGCGTACTCTCTCCGAGGCTCAGGTGGAAGATTCGTAGAGGAACAATTTAGTTTAAAG 3060
QY 3061 ACTTCATCCCTCACCTATCCCTAAACGGCTGTAGATTACCAAAAACAGATTTAATTTTCAAT 3120
DB 3061 ACTTCATCCCTCACCTATCCCTAAACGGCTGTAGATTACCAAAAACAGATTTAATTTTCAAT 3120
QY 3121 CACTAAAAGAAAAATCTATTATCAACTGCTGCTTACCAGACTTTTCTCTAGAAGCCGCTCT 3180
DB 3121 CACTAAAAGAAAAATCTATTATCAACTGCTGCTTACCAGACTTTTCTCTAGAAGCCGCTCT 3180
QY 3181 GCCTTTTACTCTGTTTTTCAAGGGACTTTTGTAAAAATCAAAATCATCTGTCAAGGCGAG 3240
DB 3181 GCCTTTTACTCTGTTTTTCAAGGGACTTTTGTAAAAATCAAAATCATCTGTCAAGGCGAG 3240
QY 3241 GAGGAGCTGATAATGAACCTTTTATTTGGAGCATTTGATCTGCAATCCAAAGGCTTCTCAGGCGG 3300
DB 3241 GAGGAGCTGATAATGAACCTTTTATTTGGAGCATTTGATCTGCAATCCAAAGGCTTCTCAGGCGG 3300
QY 3301 GCTTTGAGTGAATTTGTGTACCTGAAGTACAGTATATTTCTGTGTAATACATAAAACAAAAGC 3360
DB 3301 GCTTTGAGTGAATTTGTGTACCTGAAGTACAGTATATTTCTGTGTAATACATAAAACAAAAGC 3360
QY 3361 ATTTTGTCTAAGAGAAAGCTAATATGATTTTTTAAAGTCTAAGTCTAATATAATATGTAAG 3420
DB 3361 ATTTTGTCTAAGAGAAAGCTAATATGATTTTTTAAAGTCTAAGTCTAATATAATATGTAAG 3420
QY 3421 TTTTTCAGCTATTTAGTGATATATTTTATGGGTGGGAATAAAAATTTCTCTACAGAAAAA 3480

Db	3421	TTTTTCAGCTATT	PAGTGATATATTTTAT	TGGTGGGAATAAAATTTCTACTACAGAAAAA	3481
Qy	3481	AAAAAAAAAAAAAAAAAAAAA	3501		
Db	3481	AAAAAAAAAAAAAAAAAAAAA	3501		
RESULT 5					
AX695522					
LOCUS	AX695522	3475 bp	DNA	linear	PAT 31-MAR-2003
DEFINITION	Sequence 1149 from Patent WO03008583.				
ACCESSION	AX695522				
VERSION	AX695522.1	GI:29418674			
KEYWORDS					
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE					
AUTHORS	Morris,D.W. and Engelhard,E.K.				
TITLE	Novel compositions and methods for cancer				
JOURNAL	Patent: WO 03008583-A 1149 30-JAN-2003;				
	Sagres Discovery (US)				
FEATURES					
source	Location/Qualifiers				
	1..3475				
	/organism="Homo sapiens"				
	/mol_type="genomic DNA"				
	/db_xref="taxon:9606"				
BASE COUNT	1042 a	709 c	784 g	940 t	
ORIGIN					
Query Match	99.3%;	Score 3475;	DB 6;	Length 3475;	
Best Local Similarity	100.0%;	Pred. No. 0;			
Matches 3475;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;	
Qy	1	CGAGGCGGCAT	CGAGGCGGCTGGG	CCGCGCCCTCGGGGAC	CCCCGGGCTTCGGAGGCCATG 60
Db	1	CGAGGCGGCAT	CGAGGCGGCTGGG	CCGCGCCCTCGGGGAC	CCCCGGGCTTCGGAGGCCATG 60
Qy	61	CCGCGCTTGGCGCGGAC	CGGCGGCGCGCTGCTGCTGTTGTTTCTGCAATGATA	120	
Db	61	CCGCGCTTGGCGCGGAC	CGGCGGCGCGCTGCTGCTGTTGTTTCTGCAATGATA	120	
Qy	121	TTTGGGACTATTACAAAT	CAAGATCTGCCTCTGATCAAGTGTGTTTAAATCAATCATATAAG	180	
Db	121	TTTGGGACTATTACAAAT	CAAGATCTGCCTCTGATCAAGTGTGTTTAAATCAATCATATAAG	180	
Qy	181	AACAAATGATT	CATCAGTGGGGAAGTCATCATCATATCCATGGTATCAGAAATCCCCGGAA	240	
Db	181	AACAAATGATT	CATCAGTGGGGAAGTCATCATCATATCCATGGTATCAGAAATCCCCGGAA	240	
Qy	241	GACCTCGGGTGTGGTT	TGAGACCCAGAGCTCAGGGACAGTGACGAGCTGCGCTGTG	300	
Db	241	GACCTCGGGTGTGGTT	TGAGACCCAGAGCTCAGGGACAGTGACGAGCTGCGCTGTG	300	
Qy	301	GAAGTCAGATGATCTGCTT	CCATCACACTGCAAGTGTGTCGATGCCACAGGGAACATT	360	
Db	301	GAAGTCAGATGATCTGCTT	CCATCACACTGCAAGTGTGTCGATGCCACAGGGAACATT	360	
Qy	361	TCCTGTCTCGGGTCTTT	TAAGCAGAGCTCTCTGAATTGCCAGCACATTTTGAATTTCACAA	420	
Db	361	TCCTGTCTCGGGTCTTT	TAAGCAGAGCTCTCTGAATTGCCAGCACATTTTGAATTTCACAA	420	
Qy	421	AACAGAGGAGTGTGTT	CCATGGTCAATTTGAAATGACAGAAACCCAGCTGGAGATAC	480	
Db	421	AACAGAGGAGTGTGTT	CCATGGTCAATTTGAAATGACAGAAACCCAGCTGGAGATAC	480	
Qy	481	CTACTTTTATTTCAGAGTGAAGCTAC	CAATTTACAAATTTGTTTACAGTCAGTATAGA	540	
Db	481	CTACTTTTATTTCAGAGTGAAGCTAC	CAATTTACAAATTTGTTTACAGTCAGTATAGA	540	
Qy	541	AATACCCCTGCTTTACATTAAG	AAGACCTTACTTTAGAAAAATGGAAAAACAGGACGCC	600	


```
Qy 1681 ATCTCATTTATGCAACAAATGGTGGTTGTGTCCTCTCTTCATGTCGTTTAAACCTGCTA 1740
Db 1681 ATCTCATTTATGCAACAAATGGTGGTTGTGTCCTCTCTTCATGTCGTTTAAACCTGCTA 1740
Qy 1741 ATTGTGCACAGTACAAAGACCAATTTAGGTATGAAAGCCAGCTACAGATGGTACAGGTG 1800
Db 1741 ATTGTGCACAGTACAAAGACCAATTTAGGTATGAAAGCCAGCTACAGATGGTACAGGTG 1800
Qy 1801 ACCGGCTCCTCAGATAATAGTACTTCTACCTGTAATTTTCAGAGAAATATGAATATGATCTC 1860
Db 1801 ACCGGCTCCTCAGATAATAGTACTTCTACCTGTAATTTTCAGAGAAATATGAATATGATCTC 1860
Qy 1861 AATGGAGGTTTCCAGAGAAAATTTAGAGTTTGGAGGTTAGTACAGATCAGGTGCTTTTT 1920
Db 1861 AATGGAGGTTTCCAGAGAAAATTTAGAGTTTGGAGGTTAGTACAGATCAGGTGCTTTTT 1920
Qy 1921 GGAAGAGTGATGAACGCAACAGCTTATGGAATTAAGCAAAACAGAGAGTCTCAATCCAGGTT 1980
Db 1921 GGAAGAGTGATGAACGCAACAGCTTATGGAATTAAGCAAAACAGAGAGTCTCAATCCAGGTT 1980
Qy 1981 GCGCTCAAAATGCTGAAAGAAAAGCAGACAGCTCTGAAAGAGAGGCACTCATGTGCAAA 2040
Db 1981 GCGCTCAAAATGCTGAAAGAAAAGCAGACAGCTCTGAAAGAGAGGCACTCATGTGCAAA 2040
Qy 2041 CTCAAGATGATGACCCAGCTGGGAAGCCACGAGAATATTTGGAACCTGCTGGGGCGGTGC 2100
Db 2041 CTCAAGATGATGACCCAGCTGGGAAGCCACGAGAATATTTGGAACCTGCTGGGGCGGTGC 2100
Qy 2101 ACAGTGTGAGGACCAATTTACTGATTTTGAATTTGAACTGCTATGTTGATCTTCTCAAC 2160
Db 2101 ACAGTGTGAGGACCAATTTACTGATTTTGAATTTGAACTGCTATGTTGATCTTCTCAAC 2160
Qy 2161 TATCTAAGAGTAAAGAGAAAATTTTCAAGGACTTGGACAGAGATTTTCAAGGAACAC 2220
Db 2161 TATCTAAGAGTAAAGAGAAAATTTTCAAGGACTTGGACAGAGATTTTCAAGGAACAC 2220
Qy 2221 AATTTCAGTATTTTACCACATTTTCAATCAATCCAAATTCAGATGCTGGTGTCAAGA 2280
Db 2221 AATTTCAGTATTTTACCACATTTTCAATCAATCCAAATTCAGATGCTGGTGTCAAGA 2280
Qy 2281 GAAGTTCAGATACACCGGACTCGGATCAATCTCAGGGCTTCATGGGAATTCATTTCCAC 2340
Db 2281 GAAGTTCAGATACACCGGACTCGGATCAATCTCAGGGCTTCATGGGAATTCATTTCCAC 2340
Qy 2341 TCTGAAGATGAAATGGAATGAAAACCAAAAAGGCTGGGAAGAGAGAGGACTTGAAT 2400
Db 2341 TCTGAAGATGAAATGGAATGAAAACCAAAAAGGCTGGGAAGAGAGAGGACTTGAAT 2400
Qy 2401 GTGCTTACATTTGAAGATCTTTCTTTGCTTGCATATCAAGTTGGCCAAAGGAATGGAATTT 2460
Db 2401 GTGCTTACATTTGAAGATCTTTCTTTGCTTGCATATCAAGTTGGCCAAAGGAATGGAATTT 2460
Qy 2461 CTGGAATTTAAGTGTGTTTTCAGAGACCTGGCGCGCAGGAAGCGTCTTGTCAACCCAC 2520
Db 2461 CTGGAATTTAAGTGTGTTTTCAGAGACCTGGCGCGCAGGAAGCGTCTTGTCAACCCAC 2520
Qy 2521 GGGAAAGTGGTGAAGATATGTGACTTTTGGATTTGGCTCGAGATATCATGAGTGATTTCCAAAC 2580
Db 2521 GGGAAAGTGGTGAAGATATGTGACTTTTGGATTTGGCTCGAGATATCATGAGTGATTTCCAAAC 2580
Qy 2581 TATGTTGTGAGGGCAATGCCCCGTCTGCTGTGTAATGATGGCCCCCGAAAGCGTGTGTT 2640
Db 2581 TATGTTGTGAGGGCAATGCCCCGTCTGCTGTGTAATGATGGCCCCCGAAAGCGTGTGTT 2640
Qy 2641 GAAGGCATCTACACCATTAAGAGTGATGTCTGGTTCATATGAATATTAATCTGAGGAAATC 2700
Db 2641 GAAGGCATCTACACCATTAAGAGTGATGTCTGGTTCATATGAATATTAATCTGAGGAAATC 2700
Qy 2701 TTCTCACTTGTGTGGAATCTTACCCTGGCATTCGGTTGATGCTTAACCTTCTACAAAACCTG 2760
Db 2701 TTCTCACTTGTGTGGAATCTTACCCTGGCATTCGGTTGATGCTTAACCTTCTACAAAACCTG 2760
```

```
Qy 2761 ATTCAAATCGATTTTAAATGGATCAGCCATTTTATGCTACAGAGAAATATACATTATA 2820
Db 2761 ATTCAAATCGATTTTAAATGGATCAGCCATTTTATGCTACAGAGAAATATACATTATA 2820
Qy 2821 ATGCAATCTCTGCTGGGCTTTTGAATCAAGGAAACGGCCATCTTCCCTAAATTTGACTTCG 2880
Db 2821 ATGCAATCTCTGCTGGGCTTTTGAATCAAGGAAACGGCCATCTTCCCTAAATTTGACTTCG 2880
Qy 2881 TTTTTCAGATGTGATGCTGGCAGATGCAAGAGCGATGATACAGATGTGGATGGCCGT 2940
Db 2881 TTTTTCAGATGTGATGCTGGCAGATGCAAGAGCGATGATACAGATGTGGATGGCCGT 2940
Qy 2941 GTTTCGGAATGCTCTCACACCTACCAAAACAGGCGACCTTTTCAGCAGAGATGGAATTTG 3000
Db 2941 GTTTCGGAATGCTCTCACACCTACCAAAACAGGCGACCTTTTCAGCAGAGATGGAATTTG 3000
Qy 3001 GGGCTACTCTCTCCGAGGCTCAGGTCGAAAGATTCGTAGAGGAAACAAATTTAGTTTAAAG 3060
Db 3001 GGGCTACTCTCTCCGAGGCTCAGGTCGAAAGATTCGTAGAGGAAACAAATTTAGTTTAAAG 3060
Qy 3061 ACTTCATCTCTCCACCTATCCCTAACAGGCTGTAGATTACCAAAACAAAGATTAATTTTCA 3120
Db 3061 ACTTCATCTCTCCACCTATCCCTAACAGGCTGTAGATTACCAAAACAAAGATTAATTTTCA 3120
Qy 3121 CACTAAAGAAATCTATTATCACTGCTGTTCCAGAGACTTTTCTCTAGAACCGCTCT 3180
Db 3121 CACTAAAGAAATCTATTATCACTGCTGTTCCAGAGACTTTTCTCTAGAACCGCTCT 3180
Qy 3181 GCGTTTACTCTGTTTTCAAAGGAGCTTTTGTAAATCAAAATCATCTGTCAAGGCGAG 3240
Db 3181 GCGTTTACTCTGTTTTCAAAGGAGCTTTTGTAAATCAAAATCATCTGTCAAGGCGAG 3240
Qy 3241 GAGGAGCTGATAATGAATCTTTATTTGGAGCATTTGATCTGCATCCAGGCGCTTCTCAGGCG 3300
Db 3241 GAGGAGCTGATAATGAATCTTTATTTGGAGCATTTGATCTGCATCCAGGCGCTTCTCAGGCG 3300
Qy 3301 GCTTGAGTGAATGTGTACCTGGAAGTACAGTATATCTTGTAAATACATAAAACAAAAGC 3360
Db 3301 GCTTGAGTGAATGTGTACCTGGAAGTACAGTATATCTTGTAAATACATAAAACAAAAGC 3360
Qy 3361 ATTTTGTAAAGGAGAACTAATATGATTTTAAAGTCTATGTTTAAATATATATGTA 3420
Db 3361 ATTTTGTAAAGGAGAACTAATATGATTTTAAAGTCTATGTTTAAATATATATGTA 3420
Qy 3421 TTTTTCAGCTATTATGATGATATATTTTATGGTGGGAATAAATTTCTACTACAG 3475
Db 3421 TTTTTCAGCTATTATGATGATATATTTTATGGTGGGAATAAATTTCTACTACAG 3475
```

RESULT 6

144733 LOCUS 144733 3475 bp DNA linear PAT 07-OCT-1997
DEFINITION Sequence 3 from patent US 5635388.
ACCESSION 144733
VERSION 144733.1 GI:2469446

KEYWORDS

SOURCE

ORGANISM

Unclonable

Unclonable

REFERENCE

1 (bases 1 to 3475)

AUTHORS

Bennett, B.D., Broz, S.D., Matthews, W. and Zeigler, P.C.

TITLE

Agonist antibodies against the flk2/flt3 receptor and uses thereof

JOURNAL

Patent: US 5635388-A 3 03-JUN-1997;

FEATURES

Location/Qualifiers

source

1..3475

/organism="unknown"

BASE COUNT 1042 a 709 c 784 g 940 t

ORIGIN

Query Match

Best Local Similarity 99.3%; Score 3475; DB 6; Length 3475;

Matches 3475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 2161 TATCTAAGAAGTAAAGAGAAAAATTTTCACAGGACTTGGACAGAGATTTTTCAGGAACAC 2220
Qy 2221 AATTTTCAGTTTTTACCCCACTTTTCAATCACATCCAAATTTCCAGCATGCTCGTTCAAGA 2280
Db 2221 AATTTTCAGTTTTTACCCCACTTTTCAATCACATCCAAATTTCCAGCATGCTCGTTCAAGA 2280
Qy 2281 GAAAGTTTCAGATACACCCCGGACTCGGATCAAAATCTCAGGGCTTTTCATGGGAATTCATTTCC 2340
Db 2281 GAAAGTTTCAGATACACCCCGGACTCGGATCAAAATCTCAGGGCTTTTCATGGGAATTCATTTCC 2340
Qy 2341 TCTGAAGATGAATTAATATGATAAACCACAAAGAGCTGCGAAGAGAGAGAGACTTGAAT 2400
Db 2341 TCTGAAGATGAATTAATATGATAAACCACAAAGAGCTGCGAAGAGAGAGAGACTTGAAT 2400
Qy 2401 GTGCTTACATTTGAAGATCTCTTGTCTTGCATATCAAGTTTGCCAAAGGAATGGAATTT 2460
Db 2401 GTGCTTACATTTGAAGATCTCTTGTCTTGCATATCAAGTTTGCCAAAGGAATGGAATTT 2460
Qy 2461 CTGGAATTTAAGTCGTGTGTTTCACAGAGACCTGGCGCCGAGGAACGTGCTTGTCAACCCAC 2520
Db 2461 CTGGAATTTAAGTCGTGTGTTTCACAGAGACCTGGCGCCGAGGAACGTGCTTGTCAACCCAC 2520
Qy 2521 GGGAAAGTGGTGAAGATATGATGACTTTGATTTGGCTCGAGATATCATGAGTGAATTCACAC 2580
Db 2521 GGGAAAGTGGTGAAGATATGATGACTTTGATTTGGCTCGAGATATCATGAGTGAATTCACAC 2580
Qy 2581 TATGTTTGCAGGGCAATGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2640
Db 2581 TATGTTTGCAGGGCAATGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2640
Qy 2641 GAAGGCATCTACACCATTAAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2700
Db 2641 GAAGGCATCTACACCATTAAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2700
Qy 2701 TTCTCAGTTGGTGAATCTTACCTGCGATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2760
Db 2701 TTCTCAGTTGGTGAATCTTACCTGCGATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2760
Qy 2761 ATTCAAAATGGAATTTAAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2820
Db 2761 ATTCAAAATGGAATTTAAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2820
Qy 2821 ATGCAATCTGCTGGCTTTTGAATCAAGGAACCGCCATCTCTCCCTAAATTTGACTTGG 2880
Db 2821 ATGCAATCTGCTGGCTTTTGAATCAAGGAACCGCCATCTCTCCCTAAATTTGACTTGG 2880
Qy 2881 TTTTGTAGATGTCAGCTGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2940
Db 2881 TTTTGTAGATGTCAGCTGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2940
Qy 2941 GTTTTCGGAATGCTTCAACCTTACCAACACGCGACCTTTTCCAGCAGAGATGATGATGATGATGATGATGAT 3000
Db 2941 GTTTTCGGAATGCTTCAACCTTACCAACACGCGACCTTTTCCAGCAGAGATGATGATGATGATGATGATGAT 3000
Qy 3001 GGGCTACTCTCTCCGAGCTCAGTTCGAAGATTCGTAGAGAAACAAATTTAGTTTAAAG 3060
Db 3001 GGGCTACTCTCTCCGAGCTCAGTTCGAAGATTCGTAGAGAAACAAATTTAGTTTAAAG 3060
Qy 3061 ACTTCATCTCTCCCTCACTATCCCTAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3120
Db 3061 ACTTCATCTCTCCCTCACTATCCCTAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3120
Qy 3121 CACTAAAGAAATCTATATCACTGCTGCTTACCAGACTTTTCTCAGAGCGCTCT 3180
Db 3121 CACTAAAGAAATCTATATCACTGCTGCTTACCAGACTTTTCTCAGAGCGCTCT 3180
Qy 3181 GCGTTTACTCTTGTGTTTCAAGGGACTTTTGAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 3240
Db 3181 GCGTTTACTCTTGTGTTTCAAGGGACTTTTGAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 3240
Qy 3241 GAGGAGCTGATAATGAATCTTTATGAGCATTTGATCTGATCCAGGCTTCTCAGGCGC 3300
Db 3241 GAGGAGCTGATAATGAATCTTTATGAGCATTTGATCTGATCCAGGCTTCTCAGGCGC 3300

Qy 3301 GCTTCAGTGAATTTGTGTACCTCAAGTACAGTATATTTCTTGTAAATACATAAAACAAAGC 3360
Db 3301 GCTTCAGTGAATTTGTGTACCTCAAGTACAGTATATTTCTTGTAAATACATAAAACAAAGC 3360
Qy 3361 ATTTTCTAAGAGAGCTAATATGATTTTTTAAAGTCTATGTTTTTAAATATATATATGATAA 3420
Db 3361 ATTTTCTAAGAGAGCTAATATGATTTTTTAAAGTCTATGTTTTTAAATATATATATGATAA 3420
Qy 3421 TTTTTCAGCTATTTAGTATATATTTTATGCGTGGGAATAAAATTTCTACTACAG 3475
Db 3421 TTTTTCAGCTATTTAGTATATATTTTATGCGTGGGAATAAAATTTCTACTACAG 3475

RESULT 7
HSU02687
LOCUS
DEFINITION
Human growth factor receptor tyrosine kinase (STK-1) mRNA, complete cds.
ACCESSION
U02687
VERSION
U02687.1
GI:409572
KEYWORDS
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 3475)
Small, D., Levenstein, M., Kim, E., Carow, C., Amin, S., Rockwell, P., Witte, D., Burrow, C., Ratajczak, M. Z., Gewirtz, A. M. et al.
STK-1, the human homolog of Fik-2/Fit-3, is selectively expressed in CD34+ human bone marrow cells and is involved in the proliferation of early progenitor/stem cells
Proc. Natl. Acad. Sci. U.S.A. 91 (2), 459-463 (1994)
JOURNAL
MEDLINE
94119906
PUBMED
7507245
REFERENCE
2 (bases 1 to 3475)
Small, D.
Direct Submission
Submitted (25-OCT-1993) Donald Small, Oncology, Johns Hopkins University School of Medicine, 600 N. Wolfe St., Baltimore, MD 21287 USA
FEATURES
source
Location/Qualifiers
1..3475
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/map="13q12"
/cell_type="stem cell"
/tissue_type="bone marrow"
58..3039
/codon_start=1
/product="serine/threonine protein kinase"
/protein_id="AA18947.1"
/db_xref="GI:409573"
/translation="MPALADGALRPQSSGTVEAAAEVDVSASITLOVLDPAGNIS
SSVKGSSYPWVSPESPDLCALRPQSSGTVEAAAEVDVSASITLOVLDPAGNIS
CLWPKHSLNCQPHFLQNRGVSMVILKMTQAGVLLFIQSEATNYTLFTVSI
RNTLLTUTRRPFRKMNQDLVCISEVPPIVWVLCDSQGESCKEESPAVKKBE
KVLHFLGTDIRCCARNLQRECTRLFTIDNLQPTTLPQLFKVGEPLWIRCKAVH
VHNGFGLWLELNKALEENYPMSTYNTNRMTIRILFAFSSVARNNDGYTCSSSK
HPSOSALVTIVKGFINATNGSEYDIQYEFECFVRFKAYPQIRCTWTFGRKSPC
EOKGLDNGYSISKFCNKHQPEYIPIHAENDDAOFTKMTLNIIRKPOVLAESAQA
SCFSDGYPPLSWTHKKCDSPNCTEETEGVWNRKAKRVFGQVSSSTLWNSAIK
GFLVKCAINSIGTSCETILLNSPGFPFIQDNISFIATYIGVCLLLFIVLTLICHY
KNAYFGISKTQVSIQVAKMLKADSSREALMSELKMTQLGSHENIVNLGACT
LSPGYLTIFYCCYGDLLNLSKREKPHRTWTEIFKEHNSFYFTFQSHPNSSMPGS
REVQIHPDSDQISGLHNSFHSDEIEYENOKBLEEEDLVLTEDLLCFAYQVAKG
MEFLFKSCVRDLAARNLVTHGKVKICDFGLARDIMSDNSNVYVRGNARLPVKMA
PESLFEGLYITTKSDWVSIGILLWEIFSGVNPPIGIPVDANFYKLIQNGFKMDQPIYA
TEEYIINQSCWAFDSRKRPSFNLTLFLGCLQADAEAEAMYQNVDRVSECPHTYQNR
RPFREMDLGLLSPOAQVEDS"
BASE COUNT 1042 a 709 c 784 g 940 t

ORIGIN	Query Match Best Local Similarity Matches 3475; Conservative	99.3%; Score 3475; DB 9; Length 3475; 100.0%; Pred. No. 0; 0; Mismatches 0; Indels 0; Gaps 0;
Qy	1	CGAGGCGGCATCCGAGGGCTGGGCGGGCGCCCTGGGGACCCCGGGCTCCGAGGCCATG 60
Db	1	CGAGGCGGCATCCGAGGGCTGGGCGGGCGCCCTGGGGACCCCGGGCTCCGAGGCCATG 60
Qy	61	CCGCGGTTGGCGCGGACGCGGGCACCGTGCCGCTGCTCGTGTGTTTTCTGCAATGATA 120
Db	61	CCGCGGTTGGCGCGGACGCGGGCACCGTGCCGCTGCTCGTGTGTTTTCTGCAATGATA 120
Qy	121	TTTGGGACTATTACAAATCAAGATCTGCTGTGATCAAGTGTGTTTTTAATCAATCAATAAG 180
Db	121	TTTGGGACTATTACAAATCAAGATCTGCTGTGATCAAGTGTGTTTTTAATCAATCAATAAG 180
Qy	181	AACAAATGATTCAATCAGTGGGAAATCATCATATATCCATPGTATCAGAAATCCCGGAA 240
Db	181	AACAAATGATTCAATCAGTGGGAAATCATCATATATCCATPGTATCAGAAATCCCGGAA 240
Qy	241	GACCTCGGCTGCGGTTGAGACCCAGAGCTCAGGGACAGTGTAAGAGCTGCCGCTGTG 300
Db	241	GACCTCGGCTGCGGTTGAGACCCAGAGCTCAGGGACAGTGTAAGAGCTGCCGCTGTG 300
Qy	301	GAAGTGATGTATCTGCTTCCATCACACTGCAAGTGTGTCGATGCCACAGGGAACATT 360
Db	301	GAAGTGATGTATCTGCTTCCATCACACTGCAAGTGTGTCGATGCCACAGGGAACATT 360
Qy	361	TCCTGCTCTCGGCTTTTAAAGCACAGCTCCCTGAAATGCCAGCCACATTTTGATTACAA 420
Db	361	TCCTGCTCTCGGCTTTTAAAGCACAGCTCCCTGAAATGCCAGCCACATTTTGATTACAA 420
Qy	421	AACAGAGGAGTGTGTTTCATGGTCAATTTGAAAATGACAGAAACCCAGAGTGGAGATAC 480
Db	421	AACAGAGGAGTGTGTTTCATGGTCAATTTGAAAATGACAGAAACCCAGAGTGGAGATAC 480
Qy	481	CTACTTTTTATTACAGGTGAAGCTACCAATTPACACAATATTGTTTTACAGTGAGTAAAGA 540
Db	481	CTACTTTTTATTACAGGTGAAGCTACCAATTPACACAATATTGTTTTACAGTGAGTAAAGA 540
Qy	541	AATAACCTGCTTTACACATTAAGAAGACCTTTACTTTAGAAAAATGAAAAACCCAGACGCC 600
Db	541	AATAACCTGCTTTACACATTAAGAAGACCTTTACTTTAGAAAAATGAAAAACCCAGACGCC 600
Qy	601	CTGGTCTGCATATCTGAGAGGTTCCAGAGCCGATCGTGGAAATGGGTGCTTTTGGCATCA 660
Db	601	CTGGTCTGCATATCTGAGAGGTTCCAGAGCCGATCGTGGAAATGGGTGCTTTTGGCATCA 660
Qy	661	CAGGGGAAAGCTGTAAGAAGAAGTCCAGCTGCTGTTAAAAAGGAGGAAAAAGTCCTT 720
Db	661	CAGGGGAAAGCTGTAAGAAGAAGTCCAGCTGCTGTTAAAAAGGAGGAAAAAGTCCTT 720
Qy	721	CATCAATTTATTTGGGACGGACATTAAGGTGCTGTGCCAGAAATGAATCGGCGAGGAAATGC 780
Db	721	CATCAATTTATTTGGGACGGACATTAAGGTGCTGTGCCAGAAATGAATCGGCGAGGAAATGC 780
Qy	781	ACAGGCTGTTCAAAATAGATCTTAAATCAAACTCCTCAGACCAATTTGCCCAATTTATTT 840
Db	781	ACAGGCTGTTCAAAATAGATCTTAAATCAAACTCCTCAGACCAATTTGCCCAATTTATTT 840
Qy	841	CTTAAAGTACGGGAAACCTTATGATTAAGGTGCAAGCTGTTCAATGTGAACCAATGATTC 900
Db	841	CTTAAAGTACGGGAAACCTTATGATTAAGGTGCAAGCTGTTCAATGTGAACCAATGATTC 900
Qy	901	GGGCTCACCTGGGAATTTAGAAAACAAAGCACTCCAGGAGGGCACTACTTTGAGATGAGT 960
Db	901	GGGCTCACCTGGGAATTTAGAAAACAAAGCACTCCAGGAGGGCACTACTTTGAGATGAGT 960
Qy	961	ACCTATTTCAAACAAAGAACTATGATACGGATTCGTGTTGCTTTTGTATCATCAGTGGCA 1020
Db	961	ACCTATTTCAAACAAAGAACTATGATACGGATTCGTGTTGCTTTTGTATCATCAGTGGCA 1020

Qy	1021	AGAAACGACACCGGATACHTACACTTGTTCCTCTTCAAAGCATCCCAAGTCAATCAGCTTTG	1080
Db	1021	AGAAACGACACCGGATACHTACACTTGTTCCTCTTCAAAGCATCCCAAGTCAATCAGCTTTG	1080
Qy	1081	GTTTACCATCTGTAGGAAGGATTTATAAATGCTTACCAATTCAGTGAAGATTAATGAATTT	1140
Db	1081	GTTTACCATCTGTAGGAAGGATTTATAAATGCTTACCAATTCAGTGAAGATTAATGAATTT	1140
Qy	1141	GACCAATATGAAGAGTTTGTCTTCTGTGTCAGGTTTAAAGCCCTACCCACAAATCAGATGT	1200
Db	1141	GACCAATATGAAGAGTTTGTCTTCTGTGTCAGGTTTAAAGCCCTACCCACAAATCAGATGT	1200
Qy	1201	ACGTGGACCTTCTCTCGAAATCATTTTCCCTGTGAGCAAAAGGCTTTGATAACGGATAC	1260
Db	1201	ACGTGGACCTTCTCTCGAAATCATTTTCCCTGTGAGCAAAAGGCTTTGATAACGGATAC	1260
Qy	1261	AGCATATCCAGTTTTCGAAATCATTTTCCCTGTGAGCAAAAGGCTTTGATAACGGATAC	1320
Db	1261	AGCATATCCAGTTTTCGAAATCATTTTCCCTGTGAGCAAAAGGCTTTGATAACGGATAC	1320
Qy	1321	AATGATGATGCCCAATTTTACCAAAATGTTTCAAGCTGAAATATAAGAAGGAAACCTCAAGTG	1380
Db	1321	AATGATGATGCCCAATTTTACCAAAATGTTTCAAGCTGAAATATAAGAAGGAAACCTCAAGTG	1380
Qy	1381	CTCGAAGACATCGGCAAGTCAGGCGTCTGTTTCTCGGATGGATACCCATTTACCATCT	1440
Db	1381	CTCGAAGACATCGGCAAGTCAGGCGTCTGTTTCTCGGATGGATACCCATTTACCATCT	1440
Qy	1441	TGACCTGGAAGAGTGTTCAGACAAGTCTCCCACTGCAACAGAGATTCACAGAAGGA	1500
Db	1441	TGACCTGGAAGAGTGTTCAGACAAGTCTCCCACTGCAACAGAGATTCACAGAAGGA	1500
Qy	1501	GTCTGGAATAGAAGGCTTAACAGAAAGTGTTTGGACAGTGGGTGTCGAGCAGTACTCTA	1560
Db	1501	GTCTGGAATAGAAGGCTTAACAGAAAGTGTTTGGACAGTGGGTGTCGAGCAGTACTCTA	1560
Qy	1561	AACATGAGTGAAGCCATATAAAGGTTTCTGGTCAAAGTGTGTGCATACAATTCCTTTGGC	1620
Db	1561	AACATGAGTGAAGCCATATAAAGGTTTCTGGTCAAAGTGTGTGCATACAATTCCTTTGGC	1620
Qy	1621	ACATCTTGTGAGACGATCTCTTTTAAACTCTCTCAGGCGCTTCCCTTTCAATCCAAACAAC	1680
Db	1621	ACATCTTGTGAGACGATCTCTTTTAAACTCTCTCAGGCGCTTCCCTTTCAATCCAAACAAC	1680
Qy	1681	ATCTCATCTTATCGACAATTTGGTGTTTCTCTCTTCTCATTTGTCGTTTAAACCTTGCTA	1740
Db	1681	ATCTCATCTTATCGACAATTTGGTGTTTCTCTCTTCTCATTTGTCGTTTAAACCTTGCTA	1740
Qy	1741	ATTTGTGCACAAATGTAAGGCAATTTAGGTATGAAGCCAGCTACAGATGGTACAGGTG	1800
Db	1741	ATTTGTGCACAAATGTAAGGCAATTTAGGTATGAAGCCAGCTACAGATGGTACAGGTG	1800
Qy	1801	ACCGGCTCTCAGATAATGAGTACTTCTACGTTGATTTTCAGAGAATATGAATATGATCTC	1860
Db	1801	ACCGGCTCTCAGATAATGAGTACTTCTACGTTGATTTTCAGAGAATATGAATATGATCTC	1860
Qy	1861	AAATGGGAGTTTCCAGAGAAATTTTAGAGTTTGGGAAGGCTACAGATCAGGTGCTTTT	1920
Db	1861	AAATGGGAGTTTCCAGAGAAATTTTAGAGTTTGGGAAGGCTACAGATCAGGTGCTTTT	1920
Qy	1921	GGAAAAGTATGAACCGCAACAGCTTATGAAATTTAGCAAAACAGGAGTCTCAATCCAGGTT	1980
Db	1921	GGAAAAGTATGAACCGCAACAGCTTATGAAATTTAGCAAAACAGGAGTCTCAATCCAGGTT	1980
Qy	1981	GCCGTCAAAATGCTGAAAGAAAAGCAGACAGCTCTGAAAAGAGGSCATCTATGTCAAGAA	2040
Db	1981	GCCGTCAAAATGCTGAAAGAAAAGCAGACAGCTCTGAAAAGAGGSCATCTATGTCAAGAA	2040
Qy	2041	CTCAAGATGATGACCCAGTGGGAAGCCACGAGATATTTGTGAACTCTGTGGGGGGCGTGC	2100
Db	2041	CTCAAGATGATGACCCAGTGGGAAGCCACGAGATATTTGTGAACTCTGTGGGGGGCGTGC	2100

QY 2101 A C A C T G T C A G G A C C A A T T T A C T T G A T T T T T G A T A C T G T T G C T A T G T G A T C T T C T C A A C 2160
DB 2101 A C A C T G T C A G G A C C A A T T T A C T T G A T T T T T G A A T A C T G T T G C T A T G T G A T C T T C T C A A C 2160
QY 2161 T A T C T A A G A A G T A A A A G A G A A A A T T T C A C A G A C T T T G A C A G A G A T T T T C A A G G A A C A C 2220
DB 2161 T A T C T A A G A A G T A A A A G A G A A A A T T T C A C A G A C T T T G A C A G A G A T T T T C A A G G A A C A C 2220
QY 2221 A A T T T C A G T T T T T A C C C C A C T T T C C A A T C A C A T C C A A A T T C C A G C A T G C T G T T C A A G A 2280
DB 2221 A A T T T C A G T T T T T A C C C C A C T T T C C A A T C A C A T C C A A A T T C C A G C A T G C T G T T C A A G A 2280
QY 2281 G A A G T T C A G A T A C A C C G G A C T C G G A T C A A A T C T C A G G C T T C A T G G A A T T C A T T C A C 2340
DB 2281 G A A G T T C A G A T A C A C C G G A C T C G G A T C A A A T C T C A G G C T T C A T G G A A T T C A T T C A C 2340
QY 2341 T C T G A A G A T G A A T T G A A T A T A A A C C A A A A A A A G C T G G A A G A G A G A G A C T T G A A T 2400
DB 2341 T C T G A A G A T G A A T T G A A T A T A A A C C A A A A A A A G C T G G A A G A G A G A G A C T T G A A T 2400
QY 2401 G T G C T T A C A G A T C T T C T T T G C T T T G C A T A T C A A G T T G C C A A A G A A T G G A A T T T 2460
DB 2401 G T G C T T A C A G A T C T T C T T T G C T T T G C A T A T C A A G T T G C C A A A G A A T G G A A T T T 2460
QY 2461 C T G G A A T T T A A G T C G T G T T C A C A G A C T G G C C C C A G A A C G T G C T T G T C A C C C A C 2520
DB 2461 C T G G A A T T T A A G T C G T G T T C A C A G A C T G G C C C C A G A A C G T G C T T G T C A C C C A C 2520
QY 2521 G G G A A G T G G A G A T A T G T G A C T T T G G A T T G G C T G A T A T C A T G A T G A T T C C A C 2580
DB 2521 G G G A A G T G G A G A T A T G T G A C T T T G G A T T G G C T G A T A T C A T G A T G A T T C C A C 2580
QY 2581 T A T G T T G T C A G G G C A A T G C C C T G T G C C T G T A A A A T G G A T G G C C C C C A A A G C C T G T T 2640
DB 2581 T A T G T T G T C A G G G C A A T G C C C T G T G C C T G T A A A T G G A T G G C C C C C A A A G C C T G T T 2640
QY 2641 G A A G C A T C T A C A C C A T T A A G A T G A T G T C T G T C A T A T G A A T A T A C T G T G G G A A A T C 2700
DB 2641 G A A G C A T C T A C A C C A T T A A G A T G A T G T C T G T C A T A T G A A T A T A C T G T G G G A A A T C 2700
QY 2701 T T C T C A C T T G T G T G A T C C T T A C C T G G C A T T C C G G T G A T C T A C T T C T A C A A C T G 2760
DB 2701 T T C T C A C T T G T G A T C C T T A C C T G G C A T T C C G G T G A T C T A C T T C T A C A A C T G 2760
QY 2761 A T T C A A A A T G G A T T A A A A T G G A T C A G C C A T T T T A T G T C A G A A G A A A T A T A C A T T A T A 2820
DB 2761 A T T C A A A A T G G A T T A A A A T G G A T C A G C C A T T T T A T G T C A G A A G A A A T A T A C A T T A T A 2820
QY 2821 A T G C A A T C C T G T G G G C T T T T G A C T A A G A A A C G G C C A T C C T T C C T A A T T T G A C T T G C 2880
DB 2821 A T G C A A T C C T G T G G G C T T T T G A C T A A G A A A C G G C C A T C C T T C C C T A A T T T G A C T T G C 2880
QY 2881 T T T T T A G G A T G T C A G T G G C A G A T G C A G A A G C G A T G A T C A G A T G T G G A T G G C C G T 2940
DB 2881 T T T T T A G G A T G T C A G T G G C A G A T G C A G A A G C G A T G A T C A G A T G T G G A T G G C C G T 2940
QY 2941 G T T T C G A A T G T C T C A C A C T A C C A A A C A G G C G A C T T T C A G C A G A G A T G G A T T T G 3000
DB 2941 G T T T C G A A T G T C T C A C A C T A C C A A A C A G G C G A C T T T C A G C A G A G A T G G A T T T G 3000
QY 3001 G G G C T A C T C T C C G C A G G C T C A G G T C G A A G A T T C G T A G A G A A C A A T T A G T T T T A A G G 3060
DB 3001 G G G C T A C T C T C C G C A G G C T C A G G T C G A A G A T T C G T A G A G A A C A A T T A G T T T T A A G G 3060
QY 3061 A C T T C A T C C C T C A C C T A T C C T T A C A G G C T G T A G A T T A C C A A A C A A G A T T A A T T T C A T 3120
DB 3061 A C T T C A T C C C T C A C C T A T C C T T A C A G G C T G T A G A T T A C C A A A C A A G A T T A A T T T C A T 3120
QY 3121 C A C T A A A A G A A A T C T A T A T A C A C T C T G C T T C A C C A G A C T T T T C T A G A G C C G T C T 3180
DB 3121 C A C T A A A A G A A A T C T A T A T A C A C T C T G C T T C A C C A G A C T T T T C T A G A G C C G T C T 3180
QY 3181 . G C G T T T A C T C T T G T T T T C A A A G G A C T T T T T G T A A A A T C A A A T C A T C C T G T C A C A A G G C A G 3240

DB 3181 G G G T T T A C T C T T G T T T T C A A A G G A C T T T T T G T A A A A T C A A A T C A T C T G T C A C A G G C A G 3240
QY 3241 G A G G A C T C A T A A T G A A C T T T T A T T G G A G C A T T G A T C T G C A T C C A A G G C C T T C T C A G G C C G 3300
DB 3241 G A G G A C T C A T A A T G A A C T T T A T T G G A G C A T T G A T C T G C A T C C A A G G C C T T C T C A G G C C G 3300
QY 3301 G C T T G A G T G A A T T G T G T A C C T G A A G T A C A G A T A T A T T C T T G T A A A T A C A T A A A A A A G C 3360
DB 3301 G C T T G A G T G A A T T G T G T A C C T G A A G T A C A G A T A T A T T C T T G T A A A T A C A T A A A A A A G C 3360
QY 3361 A T T T T G C T A A G G A G A G C T A A T A T A T T T T A A G T C T A T G T T T T A A A A T A A T A T A T G T A A A 3420
DB 3361 A T T T T G C T A A G G A G A G C T A A T A T A T T T T A A G T C T A T G T T T T A A A A T A A T A T A T G T A A A 3420
QY 3421 T T T T T C A G C T A T T T A G T G A T A T A T T T T A T G G T G G G A A T A A A A T T T C T A C A C A G 3475
DB 3421 T T T T T C A G C T A T T A G T G A T A T A T T T T A T G G T G G G A A T A A A A T T T C T A C A C A G 3475

RESULT 8
I40109
LOCUS 140109 3476 bp DNA linear PAT 13-MAY-1997
DEFINITION Sequence 1 from patent US 5618709.
ACCESSION 140109
VERSION 140109.1 GI:2083114
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 3476)
AUTHORS Gewirtz,A.M., Small,D. and Civin,C.I.
TITLE Antisense oligonucleotides specific for STK-1 and method for
inhibiting expression of the STK-1 protein
JOURNAL Patent: US 5618709-A 1 08-APR-1997;
FEATURES
source Location/Qualifiers
1..3476
/organism="unknown"
BASE COUNT 1044 a 709 c 783 g 940 t
ORIGIN

Query Match 99.2%; Score 3474.4; DB 6; Length 3476;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3475; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGAGGGCGGCATCGAGGGCTGGGCCGCCGCGCGCCCTGGGGACCCCGGGCTCCGAGGCCATG 60
DB 1 CGAGGGCGGCATCGAGGGCTGGGCCGCCGCGCGCCCTGGGGACCCCGGGCTCCGAGGCCATG 60
QY 61 CGCGCGTTGGCGCGCAGCGGGCACCGTGCCTGCTCGTGTGTTTTCTGCAATGATA 120
DB 61 CGCGCGTTGGCGCGCAGCGGGCACCGTGCCTGCTCGTGTGTTTTCTGCAATGATA 120
QY 121 TTGCGGACTATTACAATCAAGATCTGCCTGTGATCAAGTGTGTTTAATCAATCAATAAG 180
DB 121 TTGCGGACTATTACAATCAAGATCTGCCTGTGATCAAGTGTGTTTAATCAATCAATAAG 180
QY 181 AACAAATGATTCATCAGTGGGGAAGTCATCATATATCCCATGTGTATCAGAAATCCCGGAA 240
DB 181 AACAAATGATTCATCAGTGGGGAAGTCATCATATATCCCATGTGTATCAGAAATCCCGGAA 240
QY 241 GACCTCGGCTGTGCGTTGAGACCCCGAGCTCAGGGACAGTGTACGAAGCTCGCGCTGTG 300
DB 241 GACCTCGGCTGTGCGTTGAGACCCCGAGCTCAGGGACAGTGTACGAAGCTCGCGCTGTG 300
QY 301 GAAGTGGATGTATCTGCTTCCATCACTGCAAGTGTGTGATGCCCGAGGAAAT 360
DB 301 GAAGTGGATGTATCTGCTTCCATCACTGCAAGTGTGTGATGCCCGAGGAAAT 360
QY 361 TCCTGTCTCTGGGTCTTTAAGACACAGCTCCCTCAATTCAGGACCATATTTGATTTCAA 420
DB 361 TCCTGTCTCTGGGTCTTTAAGACACAGCTCCCTCAATTCAGGACCATATTTGATTTCAA 420

QY 421 AACAGAGGAGTTGTTTCCATGGTCAATTTGAAATAGACAGAAACCCAGCTGGAGATAC 480
Db 421 AACAGAGGAGTTGTTTCCATGGTCAATTTGAAATAGACAGAAACCCAGCTGGAGATAC 480
QY 481 CTACTTTTATTTACAGAGTGAAGCTACCAATTAACAATATGTTTACAGTGAATTAAGA 540
Db 481 CTACTTTTATTTACAGAGTGAAGCTACCAATTAACAATATGTTTACAGTGAATTAAGA 540
QY 541 AATACCTCTCTTTACACATTTAAGAGACCTTACTTTTAGAATAATGGAACCCAGGAGCC 600
Db 541 AATACCTCTCTTTACACATTTAAGAGACCTTACTTTTAGAATAATGGAACCCAGGAGCC 600
QY 601 CTGGTCTGCATATCTGAGAGCGTTCCAGAGCCGATCGTGGATGGTCTTCCGATTC 660
Db 601 CTGGTCTGCATATCTGAGAGCGTTCCAGAGCCGATCGTGGATGGTCTTCCGATTC 660
QY 661 CAGGGGGAAGCTGTAAAGAGAAAGTCCAGCTGTTGTTTAAAGAGGAAAGTGCCT 720
Db 661 CAGGGGGAAGCTGTAAAGAGAAAGTCCAGCTGTTGTTTAAAGAGGAAAGTGCCT 720
QY 721 CATGAATTTATTTGGGACGGACATAGGTCTGTGCCAGAAATGAACTGGGCGAGGATGC 780
Db 721 CATGAATTTATTTGGGACGGACATAGGTCTGTGCCAGAAATGAACTGGGCGAGGATGC 780
QY 781 ACCAGGCTGTTCAATATAGATCTAATCAAACTCTCAGACCACATGTCACAAATTAAT 840
Db 781 ACCAGGCTGTTCAATATAGATCTAATCAAACTCTCAGACCACATGTCACAAATTAAT 840
QY 841 CTTAAAGTAGGGAACCTTTATGATAAGGTGCAAGCTGTTTCAATGTAACCATGGAATC 900
Db 841 CTTAAAGTAGGGAACCTTTATGATAAGGTGCAAGCTGTTTCAATGTAACCATGGAATC 900
QY 901 GGGCTCACCTGGGAATTAAGAAACAAAGCACTCGAGGAGGCAACTCTTTGAGATGAT 960
Db 901 GGGCTCACCTGGGAATTAAGAAACAAAGCACTCGAGGAGGCAACTCTTTGAGATGAT 960
QY 961 ACCTATTCAAACAGACATATGATACGGATTCGTTTGTGTTGTTTGTATCATCAGTGCA 1020
Db 961 ACCTATTCAAACAGACATATGATACGGATTCGTTTGTGTTGTTTGTATCATCAGTGCA 1020
QY 1021 AGAAACGACACCGGATACACTGTTGTTCTCTTCAAGCATCCAGTCAATCAGCTTTG 1080
Db 1021 AGAAACGACACCGGATACACTGTTGTTCTCTTCAAGCATCCAGTCAATCAGCTTTG 1080
QY 1081 GTTACCATCTGAGAAAGGATTTATAATGCTACCAATTCAGTGAGATTAATAAT 1140
Db 1081 GTTACCATCTGAGAAAGGATTTATAATGCTACCAATTCAGTGAGATTAATAAT 1140
QY 1141 GACCAATATGAAGATTTGTTTCTGTCAGGTTTAAAGCCTACCCACAAATCAGATGT 1200
Db 1141 GACCAATATGAAGATTTGTTTCTGTCAGGTTTAAAGCCTACCCACAAATCAGATGT 1200
QY 1201 ACCTGGACCTTCTCGAAATCATTTCTTGTGAGCAAAAGGCTCTTGATAACGGATAC 1260
Db 1201 ACCTGGACCTTCTCGAAATCATTTCTTGTGAGCAAAAGGCTCTTGATAACGGATAC 1260
QY 1261 AGCATATCCAAGTTTGGCAATCATAGCACCGCAGGAGGATATATATTTCCATGCGAA 1320
Db 1261 AGCATATCCAAGTTTGGCAATCATAGCACCGCAGGAGGATATATATTTCCATGCGAA 1320
QY 1321 AATGATGATGCCCAATTTACCAAAATGTTTCCAGCTGATATAGAGAAACCTCAAGTG 1380
Db 1321 AATGATGATGCCCAATTTACCAAAATGTTTCCAGCTGATATAGAGAAACCTCAAGTG 1380
QY 1381 CTCGCAAGAGCATCGCAAGTCTGGGCTGCTGTTTCTCGATGATACCCATTTACCCTCT 1440
Db 1381 CTCGCAAGAGCATCGCAAGTCTGGGCTGCTGTTTCTCGATGATACCCATTTACCCTCT 1440
QY 1441 TGGACCTGGAAGATGTTTACAGAAAGTCTCCCACTGACAGAGAGATCAGAGGA 1500
Db 1441 TGGACCTGGAAGATGTTTACAGAAAGTCTCCCACTGACAGAGAGATCAGAGGA 1500
QY 1501 GTCTGGAATAGAAAGCTTAACAGAAAAGTGTGAGACAGTGGGTGTCGAGCATCTCTA 1560

Db 1501 GTCTGGAATAGAAAGCTTAACAGAAAAGTGTGAGACAGTGGGTGTCGAGCATCTCTA 1560
QY 1561 AACATGAGTGAAGCCATAAAGGGTTCCTGGTCAAGTGTGTGATCAATTCCTTGGC 1620
Db 1561 AACATGAGTGAAGCCATAAAGGGTTCCTGGTCAAGTGTGTGATCAATTCCTTGGC 1620
QY 1621 ACATCTGTGAGACGATCCCTTTTAAACTCTCCAGGCCCTTCCCTTTTCAATCCAGACAAC 1680
Db 1621 ACATCTGTGAGACGATCCCTTTTAAACTCTCCAGGCCCTTCCCTTTTCAATCCAGACAAC 1680
QY 1681 ATCTCATTTCTATGCAACAATTTGTTGTCTCTCTTCAATTTGCTTTTAAACCTGCTA 1740
Db 1681 ATCTCATTTCTATGCAACAATTTGTTGTCTCTCTTCAATTTGCTTTTAAACCTGCTA 1740
QY 1741 ATTGTGTCACAAGTACAAAAGCAATTTAGTATGAAAGCCAGCTACAGATGTTACAGGTG 1800
Db 1741 ATTGTGTCACAAGTACAAAAGCAATTTAGTATGAAAGCCAGCTACAGATGTTACAGGTG 1800
QY 1801 ACCGGCTCTCAGATAATGAGTACTTCTACGTTGATTTTCAGAGAAATATGAATATGATCTC 1860
Db 1801 ACCGGCTCTCAGATAATGAGTACTTCTACGTTGATTTTCAGAGAAATATGAATATGATCTC 1860
QY 1861 AAATGGAGTTTCCAGAGAAAATTTAGAGTTTGGGAGGTACTAGGATCAGGTGCTTTT 1920
Db 1861 AAATGGAGTTTCCAGAGAAAATTTAGAGTTTGGGAGGTACTAGGATCAGGTGCTTTT 1920
QY 1921 GGAAAGTGATGAAACGCAACAGCTTATGGAATTTAGCAAAACAGGAGTCTCAATCCAGTT 1980
Db 1921 GGAAAGTGATGAAACGCAACAGCTTATGGAATTTAGCAAAACAGGAGTCTCAATCCAGTT 1980
QY 1981 GCCGTCAAAATGCTCAAAAGAAAAGCAGACAGCTCTGAAAGAGAGCACTCATGTGAGAA 2040
Db 1981 GCCGTCAAAATGCTGAAAGAAAAGCAGACAGCTCTGAAAGAGAGCACTCATGTGAGAA 2040
QY 2041 CTCAGATGATGACCCAGTGGGAGCCACAGAGATATTGTCAACTGCTGGGGGGCTGC 2100
Db 2041 CTCAGATGATGACCCAGTGGGAGCCACAGAGATATTGTCAACTGCTGGGGGGCTGC 2100
QY 2101 ACCTGTGAGACCAATTTACTGTTTGAATTTGCTATGCTGCTGCTCTTCAAC 2160
Db 2101 ACCTGTGAGACCAATTTACTGTTTGAATTTGCTATGCTGCTGCTTCTCAAC 2160
QY 2161 TATCTAAGAGTAAAGAGAAAATTTTCAAGGACTTGGACAGAGATTTTCAAGGAAACAC 2220
Db 2161 TATCTAAGAGTAAAGAGAAAATTTTCAAGGACTTGGACAGAGATTTTCAAGGAAACAC 2220
QY 2221 AATTTCAAGTTTACCCCACTTTTCCAATCAGATCCAAATTTCCAGATGCTTGAAT 2280
Db 2221 AATTTCAAGTTTACCCCACTTTTCCAATCAGATCCAAATTTCCAGATGCTTGAAT 2280
QY 2281 GAAGTTCAGATACACCCGACTCGGATCAATCTCAGGGCTTCATGGAAATTTCAATTCAC 2340
Db 2281 GAAGTTCAGATACACCCGACTCGGATCAATCTCAGGGCTTCATGGAAATTTCAATTCAC 2340
QY 2341 TCTGAAGATGAATTTGAATATGAAACCAAAAGGCTGGAAGAGAGGAGGACTTGAAT 2400
Db 2341 TCTGAAGATGAATTTGAATATGAAACCAAAAGGCTGGAAGAGAGGAGGACTTGAAT 2400
QY 2401 GTGCTTACATTTGAAAGATCTTTTGTGCTTATCATCAAGTTTCCAAAGAAATGGAATTT 2460
Db 2401 GTGCTTACATTTGAAAGATCTTTTGTGCTTATCATCAAGTTTCCAAAGAAATGGAATTT 2460
QY 2461 CTGGAATTTAAGTCTGTTTCAAGAGACTTGGCCGCGCAGGAAACGCTGTTGTCAACCCAC 2520
Db 2461 CTGGAATTTAAGTCTGTTTCAAGAGACTTGGCCGCGCAGGAAACGCTGTTGTCAACCCAC 2520
QY 2521 GGGAGAGTGGTCAAGATGTCATTTGGATTTGGCTCGAGATATCATGATGATTTCAAC 2580
Db 2521 GGGAGAGTGGTCAAGATGTCATTTGGATTTGGCTCGAGATATCATGATGATTTCAAC 2580
QY 2581 TATGTTGTGAGGGGCAATGCCCGTCTGCTGTAATATGATGCTGCTGCTGCTGCTGCTT 2640

QY 298 GTGGAAGTGGATGATCTGCTTCCATCACATGCAAGTGTGGTTCGATGCCCGAGGAAC 357
DB 241 GTGGAAGTGGATGATCTGCTTCCATCACATGCAAGTGTGGTTCGATGCCCGAGGAAC 300
QY 358 ATTTCTGCTCTCGGCTTTAAGCACAGCTCCCTGGAATGCGACGCCACATTTTGAATTA 417
DB 301 ATTTCTGCTCTCGGCTTTAAGCACAGCTCCCTGGAATGCGACGCCACATTTTGAATTA 360
QY 418 CAAACACAGAGAGTGTGTTCCATGTCATTTTGAATAATGACAGAAACCCAGCTGGAA 477
DB 361 CAAACACAGAGAGTGTGTTCCATGTCATTTTGAATAATGACAGAAACCCAGCTGGAA 420
QY 478 TACCTACTTTTTATTTCAGAGTGAAGCTACCAATTACAAATATTTTACAGTGTAGTATA 537
DB 421 TACCTACTTTTTATTTCAGAGTGAAGCTACCAATTACAAATATTTTACAGTGTAGTATA 480
QY 538 AGAATAACCTGCTTTACACATTAAGAAAGACCTTACTTTAGAAAAATGGAACACAGAC 597
DB 481 AGAATAACCTGCTTTACACATTAAGAAAGACCTTACTTTAGAAAAATGGAACACAGAC 540
QY 598 GCCCTGCTGCATATCTGAGAGCGTTCAGAGCGGTTCAGAGCGCATGCGGAATGGTCTTTGCCAT 657
DB 541 GCCCTGCTGCATATCTGAGAGCGTTCAGAGCGCATGCGGAATGGTCTTTGCCAT 600
QY 658 TCACAGGGGGAAGCTGTAAAGAAAGAAAGTCCAGCTGTTGTTTAAAGAGGAGAAAGTG 717
DB 601 TCACAGGGGGAAGCTGTAAAGAAAGAAAGTCCAGCTGTTGTTTAAAGAGGAGAAAGTG 660
QY 718 CTTTCATGAATTAATTTGGGACGCAATAAGTGTGTCGACAGAAATGAACTGGGACGGAA 777
DB 661 CTTTCATGAATTAATTTGGGACGCAATAAGTGTGTCGACAGAAATGAACTGGGACGGAA 720
QY 778 TGCACACAGGCTGTTCAAAATAGATCTAAATCAAACTCCTCAGACCAATGTCACAAATTA 837
DB 721 TGCACACAGGCTGTTCAAAATAGATCTAAATCAAACTCCTCAGACCAATGTCACAAATTA 780
QY 838 TTTCTTAAAGTAGGGACCTTATGATAGAGTGCAAAGCTGTTTCATGTGAACCATGA 897
DB 781 TTTCTTAAAGTAGGGAAACCTTATGATAGAGTGCAAAGCTGTTTCATGTGAACCATGA 840
QY 898 TTCGGGCTCACCTGGGAATTAGAAAAACAAAGCACTCGAGGAGGGCAACTACTTTGAGATG 957
DB 841 TTCGGGCTCACCTGGGAATTAGAAAAACAAAGCACTCGAGGAGGGCAACTACTTTGAGATG 900
QY 958 AGTACCTATTCAACAAACAGAACTATGATACGGATTCTGTTGCTTTGTTGATATCATCAGTG 1017
DB 901 AGTACCTATTCAACAAACAGAACTATGATACGGATTCTGTTGCTTTGTTGATATCATCAGTG 960
QY 1018 GCAAGAAACGACACCGGATACACTTGTTCCTCTTCAAAGCATCCAGTCAATCAGCT 1077
DB 961 GCAAGAAACGACACCGGATACACTTGTTCCTCTTCAAAGCATCCAGTCAATCAGCT 1020
QY 1078 TTGTTTACCATCGTAGGAAAGGGATTATTAATGCTACCAATTCAGGTGAAGATTATGAA 1137
DB 1021 TTGTTTACCATCGTAGAAGAGGATTATTAATGCTACCAATTCAGGTGAAGATTATGAA 1080
QY 1138 ATTGACCAATATGAAGAGTTTGTGTTTTCTGTAGGTTTAAAGCCTACCCCAAAATCAGA 1197
DB 1081 ATTGACCAATATGAAGAGTTTGTGTTTTCTGTAGGTTTAAAGCCTACCCCAAAATCAGA 1140
QY 1198 TGTACGTGGACCTCTCTCGAAATCATTTCCCTGTGAGCAAAAGGGTCTTGATTAACGGA 1257
DB 1141 TGTACGTGGACCTCTCTCGAAATCATTTCCCTGTGAGCAAAAGGGTCTTGATTAACGGA 1200
QY 1258 TACAGCATATCAAGTTTGTGCAATCATAAGCACAGCCAGGAGAGATATATATTCATGCA 1317
DB 1201 TACAGCATATCAAGTTTGTGCAATCATAAGCACAGCCAGGAGAGATATATATTCATGCA 1260
QY 1318 GAAATGATGATGCCCAATTTTACCAAAATGTTTCAGCTGAAATATGAAGAGGAAACCTCAA 1377
DB 1261 GAAATGATGATGCCCAATTTTACCAAAATGTTTCAGCTGAAATATGAAGAGGAAACCTCAA 1320
QY 1378 GTGCTCGCAGAAAGCATCGGCAAGTCAAGGCGTCTGTTCTCGGATGGATACCCATTACCA 1437

DB 1321 GTGCTCGCAGAAAGCATCGGCAAGTCAAGGCTCTGTTCTCGGATGGATACCCATTACCA 1380
QY 1438 TCTTGGACCTCGAAGAGAGTGTTCAGACAAAGTCTCCAACTGCAACAGAGAGATCACAGAA 1497
DB 1381 TCTTGGACCTCGAAGAGAGTGTTCAGACAAAGTCTCCAACTGCAACAGAGAGATCACAGAA 1440
QY 1498 GGAAGTCTGGAATAGAAAGGCTAAACAGAAAGTGTGTTGGACAGTGGGTGTCGAGCAGTACT 1557
DB 1441 GGAAGTCTGGAATAGAAAGGCTAAACAGAAAGTGTGTTGGACAGTGGGTGTCGAGCAGTACT 1500
QY 1558 CTAACATGAGTGAAGCCATAAAGGGTTCCTGGTCAAGTGTGTCATACATAAATCCCTTT 1617
DB 1501 CTAACATGAGTGAAGCCATAAAGGGTTCCTGGTCAAGTGTGTCATACATAAATCCCTTT 1560
QY 1618 GGCACATCTGTGAGACGATCCTTTTAAACTCTCCAGGCCCTTCCCTTTTCATCCAAAGAC 1677
DB 1561 GGCACATCTGTGAGACGATCCTTTTAAACTCTCCAGGCCCTTCCCTTTTCATCCAAAGAC 1620
QY 1678 AACATCTCTATGCAACAAATTTGGTGTGTTGTCCTCTTCATTTGTCGTTTAAACCTTG 1737
DB 1621 AACATCTCTATGCAACAAATTTGGTGTGTTGTCCTCTTCATTTGTCGTTTAAACCTTG 1680
QY 1738 CTAATTTGTCAAAAGTACAAAAGCAATTTAGGTATGAAAGCCAGCTACAGATGGTACAG 1797
DB 1691 CTAATTTGTCAAAAGTACAAAAGCAATTTAGGTATGAAAGCCAGCTACAGATGGTACAG 1740
QY 1798 GTGACCGGCTCCTCAGATAATGAGTACTTCTACGTTGATTTTCCAGAGAAATATGAATATGAT 1857
DB 1741 GTGACCGGCTCCTCAGATAATGAGTACTTCTACGTTGATTTTCCAGAGAAATATGAATATGAT 1800
QY 1858 CTCAAATGGGAGTTTCCAAAGAGAAATTTAGAGTTTGGAGAGTACTAGGATCAGGTGCT 1917
DB 1801 CTCAAATGGGAGTTTCCAAAGAGAAATTTAGAGTTTGGAGAGTACTAGGATCAGGTGCT 1860
QY 1918 TTTGGAAGTGTGAAACGCAACAGCTTATGGAATTTAGCAAAACAGGAGTCTCAATCCAG 1977
DB 1861 TTTGGAAGTGTGAAACGCAACAGCTTATGGAATTTAGCAAAACAGGAGTCTCAATCCAG 1920
QY 1978 GTTGCCTGCAAAATGCTGAAAGAAAAAGCAGACAGCTCTGAAAGAGAGGCACTCATGTCA 2037
DB 1921 GTTGCCTGCAAAATGCTGAAAGAAAAAGCAGACAGCTCTGAAAGAGAGGCACTCATGTCA 1980
QY 2038 GAACTCAAGATGATGACCCAGCTGGAAAGCCACAGAAATATGTAACCTGCTGGGGGG 2097
DB 1981 GAACTCAAGATGATGACCCAGCTGGAAAGCCACAGAAATATGTAACCTGCTGGGGGG 2040
QY 2098 TGCACTGTGAGGACCAATTTACTTGTATTTTGAATTTGTAATGCTGTTGCTATGGTCTCTTC 2157
DB 2041 TGCACTGTGAGGACCAATTTACTTGTATTTTGAATTTGTAATGCTGTTGCTATGGTCTCTTC 2100
QY 2158 AACTATCTAAGAGTAAAGAGAAAAATTTTCACAGGACTTGGACAGAGATTTTCAAGGAA 2217
DB 2101 AACTATCTAAGAGTAAAGAGAAAAATTTTCACAGGACTTGGACAGAGATTTTCAAGGAA 2160
QY 2218 CACAATTTTCACTTTTACCCCACTTTTCCAATCACAATCCAAATTTCCAGCATGCTGTTCA 2277
DB 2161 CACAATTTTCACTTTTACCCCACTTTTCCAATCACAATCCAAATTTCCAGCATGCTGTTCA 2220
QY 2278 AGAGAGTTTCAAGATACCCCGACTCGGATCAATCTCAGGGCTTCATGGGAATTCATTT 2337
DB 2221 AGAGAGTTTCAAGATACCCCGACTCGGATCAATCTCAGGGCTTCATGGGAATTCATTT 2280
QY 2338 CACTCTGAAGATGAAATGGAATATGAAAAACAAAAAGGCTGGAAGAGAGAGGAGACTTG 2397
DB 2281 CACTCTGAAGATGAAATGGAATATGAAAAACAAAAAGGCTGGAAGAGAGGAGACTTG 2340
QY 2398 AATGTGCTTACATTTGAAGATCTTTTGTGCTTTCATATCAAGTTGCCAAAGGAATGGAA 2457
DB 2341 AATGTGCTTACATTTGAAGATCTTTTGTGCTTTCATATCAAGTTGCCAAAGGAATGGAA 2400
QY 2458 TTTCTGGAAATTAAGTGTGTTTTCAGAGACCTTGGCGCCAGGAACTGCTTGTCTCACC 2517

D _b	781	TTTCTTAAAGTAGGGGAACCCCTTATGGAATAAGGTGCAAGCTGTTTCATTGTGAACCATCGA	840
Q _y	898	TTCGGGCTCACCTGGGAATTAGAAAACAAGCATTGAGGAGGGCAACTACTCTTTTGAGATG	957
D _b	841	TTCGGGCTCACCTGGGAATTAGAAAACAAGCATTGAGGAGGGCACTACTCTTTTGAGATG	900
Q _y	958	AGTACTCTATTCAAACAAAAAGAACTATGATACGGATTCCTTTGCTTTTGTATCATCATG	1017
D _b	901	AGTACTCTATTCAAACAAAAAGAACTATGATACGGATTCCTTTGCTTTTGTATCATCATG	960
Q _y	1018	GCAAGAAACGACACCGGATCTACACTTTGTCCTCTTCAAAGCATCCAGTCAATCATGCT	1077
D _b	961	GCAAGAAACGACACCGGATCTACACTTTGTCCTCTTCAAAGCATCCAGTCAATCATGCT	1020
Q _y	1078	TTGGTTACCATCTGAGGAAAGGGATTATAAATGCTACCAATCAAGTGAAGATTATGAA	1137
D _b	1021	TTGGTTACCATCTGAGGAAAGGGATTATAAATGCTACCAATCAAGTGAAGATTATGAA	1080
Q _y	1138	ATTGACCATAATGAAAGATTTTGTGTTTTCTGTCAAGGTTTAAGGCCTACCCAACAATCAGA	1197
D _b	1081	ATTGACCATAATGAAAGATTTTGTGTTTTCTGTCAAGGTTTAAGGCCTACCCAACAATCAGA	1140
Q _y	1198	TGTACGTGGACCTTCTCTCGAAATCATTTCCTTGTGAGCAAAAAGGCTCTTGATAACGGA	1257
D _b	1141	TGTACGTGGACCTTCTCTCGAAATCATTTCCTTGTGAGCAAAAAGGCTCTTGATAACGGA	1200
Q _y	1258	TACAGCATATCCAAGTTTGTGCAATCATAAAGCACGAGCAGAGAAATATATATCCCATGCA	1317
D _b	1201	TACAGCATATCCAAGTTTGTGCAATCATAAAGCACGAGCAGAGAAATATATATCCCATGCA	1260
Q _y	1318	GAAATGATGATGCCCAATTTACCAAAATGTTTCACGCTGMAATATAAGAAAGAAACCTCAA	1377
D _b	1261	GAAATGATGATGCCCAATTTACCAAAATGTTTCACGCTGMAATATAAGAAAGAAACCTCAA	1320
Q _y	1378	GTGCTCGCAGNAGCATCGGCAAGTCAGCGCTCTGTTTCTCGATGGATACCCATTAACCA	1437
D _b	1321	GTGCTCGCAGNAGCATCGGCAAGTCAGCGCTCTGTTTCTCGATGGATACCCATTAACCA	1380
Q _y	1438	TC TTGGACCTGGAAGAAAGTGTTCAGACAAGTCTCCCAACTGCACAGAGAGATCACAGAA	1497
D _b	1381	TC TTGGACCTGGAAGAAAGTGTTCAGACAAGTCTCCCAACTGCACAGAGAGATCACAGAA	1440
Q _y	1498	GGAGTCTGGAATAGAAAAGGCTTAAACAGAAAGTGTTTGGAACAGTGGGTGTGAGCAGTACT	1557
D _b	1441	GGAGTCTGGAATAGAAAAGGCTTAAACAGAAAGTGTTTGGAACAGTGGGTGTGAGCAGTACT	1500
Q _y	1558	CTAAACATGAGTGAAGCCATAAAGGGTTCCTGTGCAAGTCTGTGTGCATACAAATCCCTTT	1617
D _b	1501	CTAAACATGAGTGAAGCCATAAAGGGTTCCTGTGCAAGTCTGTGTGCATACAAATCCCTTT	1560
Q _y	1618	GGACATCTTTGTGAGACGATCCTTTTAAACTCTCCAGGCCCTTCCCTTTTCATCCCAAGAC	1677
D _b	1561	GGACATCTTTGTGAGACGATCCTTTTAAACTCTCCAGGCCCTTCCCTTTTCATCCCAAGAC	1620
Q _y	1678	AACATCTCATTTCTATGCAACAATATGGTGTGTTGTCTCTCTTCATTGTGCTTTTAAACCTG	1737
D _b	1621	AACATCTCATTTCTATGCAACAATATGGTGTGTTGTCTCTCTTCATTGTGCTTTTAAACCTG	1680
Q _y	1738	CTAATTTGTGCAAGTACAAAAGCAATTTAGGTATGAAAGCCAGCTPACAGATGGTACAG	1797
D _b	1681	CTAATTTGTGCAAGTACAAAAGCAATTTAGGTATGAAAGCCAGCTPACAGATGGTACAG	1740
Q _y	1798	GTGACCGGCTCCTCAGATAATGAGTACTCTACGTTGATTTTCAGAGNAATGATGATGAT	1857
D _b	1741	GTGACCGGATCCTCAGATAATGAGTACTCTACGTTGATTTTCAGAGNAATGATGATGAT	1800
Q _y	1858	CTCAAAATGGAGTTTCCAAGAGAAAAATTTAGAGTTTGGGAAGGTACTAGGATCAGGTGCT	1917
D _b	1801	GTCAAAATGGAGTTTCCAAGAGAAAAATTTAGAGTTTGGGAAGGTACTAGGATCAGGTGCT	1860
Q _y	1918	TTTGGAAAAGTGTATGAACGCAACAGCTTATGGAATTTAGCAAAAACAGGAGTCTCAATCCAG	1977
D _b	1861	TTTGGAAAAGTGTATGAACGCAACAGCTTATGGAATTTAGCAAAAACAGGAGTCTCAATCCAG	1920

Qy	3058	AGGACTTCATCCCTCCACCTTATCCCTAACAGGCTGTAGATTACCAAAACAAGATTAAATTT	3111
Db	3001	AGGACTTCATCCCTCCACCTTATCCCTAACAGGCTGTAGATTACCAAAACAAGGTTAATTT	3060
Qy	3118	CATCACTAAAGAAATCTATTATCAACTGCTGCTTCACAGAGCTTTTCTCTAGAGCGG	3177
Db	3061	CATCACTAAAGAAATCTATTATCAACTGCTGCTTCACAGAGCTTTTCTCTAGAGAGCG	3120
RESULT 11			
I44518			
LOCUS	I44518	3120 bp	DNA linear PAT 07-OCT-1997
DEFINITION	Sequence 22 from patent US 5635177.		
ACCESSION	I44518		
VERSION	I44518.1 GI:2469231		
KEYWORDS			
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	Unclassified.		
AUTHORS	1 (bases 1 to 3120)		
TITLE	Bennett, B.D., Goeddel, D. and Matthews, W.		
JOURNAL	Protein tyrosine kinase agonist antibodies		
FEATURES	Patent: US 5635177-A 22 03-JUN-1997;		
source	Location/Qualifiers		
	1..3120		
	/organism="unknown"		
BASE COUNT	944 a	642 c	700 g 834 t
ORIGIN			
Query Match 88.1%; Score 3083.2; DB 6; Length 3120;			
Best Local Similarity 99.3%; Pred. No. 0;			
Matches 3097; Conservative 0; Mismatches 23; Indels 0; Gaps 0;			
Qy	58	ATGCCGGGTGGCGCGGACCGTGCCTGATCAAGTGTGTTTAAATCAATCAT	117
Db	1	ATGAGAGCGTGGCGCGGACCGTGCCTGATCAAGTGTGTTTAAATCAATCAT	60
Qy	118	ATATTTGGGACTATTACAAATCAAGATCTGCCTGTGATCAAGTGTGTTTAAATCAATCAT	177
Db	61	ATATTTGGGACTATTACAAATCAAGATCTGCCTGTGATCAAGTGTGTTTAAATCAATCAT	120
Qy	178	AAGAAACATGATTCATCAGTGGGGAATCATCATATCCATGGTATCAGAATCCCG	237
Db	121	AAGAAACATGATTCATCAGTGGGGAATCATCATATCCATGGTATCAGAATCCCG	180
Qy	238	GAAGACCTCGGTTGGTGGAGACCCAGAGCTCAGGACAGTGTACGAGCTGCCGCT	297
Db	181	GAAGACCTCGGTTGGTGGAGACCCAGAGCTCAGGACAGTGTACGAGCTGCCGCT	240
Qy	298	GTGGAAGTGGATGATCTGCTTCCATCACACTGCAAGTGTGTCGATGCCCAAGGAAC	357
Db	241	GTGGAAGTGGATGATCTGCTTCCATCACACTGCAAGTGTGTCGATGCCCAAGGAAC	300
Qy	358	ATTTCTGTCTCTGGGCTTTAAGCACAGCTCCCTGAAATGCCAGCCACATTTTGATTTA	417
Db	301	ATTTCTGTCTCTGGGCTTTAAGCACAGCTCCCTGAAATGCCAGCCACATTTTGATTTA	360
Qy	418	CAAAACAGAGAGTGTGTTTCATGTTGAAAATGACAGAAAACCCAGCTGGAGAA	477
Db	361	CAAAACAGAGAGTGTGTTTCATGTTGAAAATGACAGAAAACCCAGCTGGAGAA	420
Qy	478	TACCTACTTTTATTCAGAGTGAAGTACCAATTACAAATTTGTTTACAGTGAGTATA	537
Db	421	TACCTACTTTTATTCAGAGTGAAGTACCAATTACAAATTTGTTTACAGTGAGTATA	480
Qy	538	AGAAATACCCCTGCTTACACATTAAGAAAGCTTACTTTAGAAAATGGAATAACAGGAC	597
Db	481	AGAAATACCCCTGCTTACACATTAAGAAAGCTTACTTTAGAAAATGGAATAACAGGAC	540
Qy	598	GCCCTGCTGCATATCTGAGAGCGTTCCAGAGCCGATCGTGGAAATGGGTGCTTTGCGAT	657
Db	541	GCCCTGCTGCATATCTGAGAGCGTTCCAGAGCCGATCGTGGAAATGGGTGCTTTGCGAT	600

```
Db 1681 CTAATTTGTCAAGTACAAAGCAATTTAGGTATGAAGCCAGCTACAGATGATACAG 1740
QY 1798 GTACACGGCTCCTCAGATATAGTACTCTACGTTGATTTTTCAGAGATATGATATGAT 1857
Db 1741 GTACCGGATTCCTCAGATATAGTACTCTACGTTGATTTTTCAGAGATATGATATGAT 1800
QY 1858 CTCAATGGGAGTTTCCAGAGAGAAATTTAGAGTTTGGGAAGGTACTAGGATCAGGTGCT 1917
Db 1801 GTCAATGGGAGTTTCCAGAGAGAAATTTAGAGTTTGGGAAGGTACTAGGATCAGGTGCT 1860
QY 1918 TTTGGAAGAGTGATGATGACGCAACAGCTTTATGGAATTTAGCAAAAACAGAGAGTCTCAATCCAG 1977
Db 1861 TTTGGAAGAGTGATGATGACGCAACAGCTTTATGGAATTTAGCAAAAACAGAGAGTCTCAATCCAG 1920
QY 1978 GTTGCCGTCAAAATGCTGAAGAGAAAGCAGACAGCTCTGAAAGAGAGGCACTCATGTCA 2037
Db 1921 GTTACCGTCAAAATGCTGAAGAGAAAGCAGACAGCTCTGAAAGAGAGGCACTCATGTCA 1980
QY 2038 GAACTCAAGATGATGACCCAGCTGGGAAGCCAGAGAAATTTGTGAACCTGCTGGGGGGG 2097
Db 1981 GAACTCAAGATGATGACCCAGCTGGGAAGCCAGAGAAATTTGTGAACCTGCTGGGGGGG 2040
QY 2098 TGCACACTGTCAGGACCAATTTACTTGAATTTTGAATCTGTTGCTATGCTGATCTTCTC 2157
Db 2041 TGCACACTGTCAGGACCAATTTACTTGAATTTTGAATCTGTTGCTATGCTGATCTTCTC 2100
QY 2158 AACTATCTAAGAGTAAAGAGAGAAATTTTCAAGACTTTGACAGAGATTTTCAAGGAA 2217
Db 2101 AACTATCTAAGAGTAAAGAGAGAAATTTTCAAGACTTTGACAGAGATTTTCAAGGAA 2160
QY 2218 CACAATTTTCAAGTAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATG 2277
Db 2161 CACAATTTTCAAGTAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATG 2220
QY 2278 AGAGAAGTTTCAAGTAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATG 2337
Db 2221 AGAGAAGTTTCAAGTAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATG 2280
QY 2338 CACTCTGAAGTAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATG 2397
Db 2281 CACTCTGAAGTAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATG 2340
QY 2398 AATGTCCTTACATTTGAGATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 2457
Db 2341 AATGTCCTTACATTTGAGATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 2400
QY 2458 TTTCTGGAATTTAAGTCGTGTTTCAAGAGACCTGGCGCCAGGAACGTGCTTGTTCACC 2517
Db 2401 TTTCTGGAATTTAAGTCGTGTTTCAAGAGACCTGGCGCCAGGAACGTGCTTGTTCACC 2460
QY 2518 CACGGGAAAGTGGTGAAGATATGACATTTGGATTTGGCTCGAGATATCATGATGATTC 2577
Db 2461 CACGGGAAAGTGGTGAAGATATGACATTTGGATTTGGCTCGAGATATCATGATGATTC 2520
QY 2578 AACTATGTTGTCAGGGCAATCCCGTCTGCTGTTGATTTGATTTGATTTGATTTGATTTGAT 2637
Db 2521 AACTATGTTGTCAGGGCAATCCCGTCTGCTGTTGATTTGATTTGATTTGATTTGATTTGAT 2580
QY 2638 TTTGAAGGCATCTACACCATTTAAGAGTATGATGCTGATTTGATTTGATTTGATTTGATTTGAT 2697
Db 2581 TTTGAAGGCATCTACACCATTTAAGAGTATGATGCTGATTTGATTTGATTTGATTTGATTTGAT 2640
QY 2698 ATCTTCTCACTTGGTGTGAATCTTACCTTGGCAATTCGGTGTGATTTGATTTGATTTGATTT 2757
Db 2641 ATCTTCTCACTTGGTGTGAATCTTACCTTGGCAATTCGGTGTGATTTGATTTGATTTGATTT 2700
QY 2758 CTGATTTCAAAATGATTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2817
Db 2701 CTGATTTCAAAATGATTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2760
QY 2818 ATATGCAATCTCTGGGCTTTTGAATTTTGAATTTTGAATTTTGAATTTTGAATTTTGAATTT 2877
```

```
Db 2761 ATAATGCAATCCTCTGGGCTTTTGAATTTTGAATTTTGAATTTTGAATTTTGAATTTTGAAT 2820
QY 2878 TCGTTTTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2937
Db 2821 TCGTTTTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2880
QY 2938 CGTGTTCGGAATGCTCTCACCTACCACTACCACTACCACTACCACTACCACTACCACTACCA 2997
Db 2881 CGTGTTCGGAATGCTCTCACCTACCACTACCACTACCACTACCACTACCACTACCACTACCA 2940
QY 2998 TTTGGGGCTACTCTCTCCGAGGCTCAGGTGCAAGATTTCGTAGAGGAAACAAATTTAGTTT 3057
Db 2941 TTTGGGGCTACTCTCTCCGAGGCTCAGGTGCAAGATTTCGTAGAGGAAACAAATTTAGTTT 3000
QY 3058 AGGACTTTCCTCTCCACCTATCCCTAACAGGCTGTAGATTACCAAAACAGATTTAAATTT 3117
Db 3001 AGGACTTTCCTCTCCACCTATCCCTAACAGGCTGTAGATTACCAAAACAGATTTAAATTT 3060
QY 3118 CATCACTTAAAGAAATCTATTATCACTGCTGCTTCCAGAGACTTTTCTTAGAAGCG 3177
Db 3061 CATCACTTAAAGAAATCTATTATCACTGCTGCTTCCAGAGACTTTTCTTAGAAGCG 3120

RESULT 12
AX695523
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1
AUTHORS
MORRIS, D.W. and ENGELHARD, E.K.
TITLE
Novel compositions and methods for cancer
JOURNAL
Patent: WO 0308583-A 1150 30-JAN-2003;
Sagres Discovery (US)
FEATURES
Location/Qualifiers
1..2982
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT
896 a 612 c 683 g 791 t
ORIGIN
Query Match 85.2%; Score 2982; DB 6; Length 2982;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2982; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 58 ATGCCGGGCTTGGCGCGGACGCGGGCACCGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 117
Db 1 ATGCCGGGCTTGGCGCGGACGCGGGCACCGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 60
QY 118 ATATTGGGACTATTACAAATCAAGATCTGCTGTGATCAAGTGTGTTTAAATCAATCAT 177
Db 61 ATATTGGGACTATTACAAATCAAGATCTGCTGTGATCAAGTGTGTTTAAATCAATCAT 120
QY 178 AAGAAATGATTTATCATAGTGGGGAAGTCAATCATATATCCATGATTCAGATATCCCG 237
Db 121 AAGAAATGATTTATCATAGTGGGGAAGTCAATCATATATCCATGATTCAGATATCCCG 180
QY 238 GAAGACCTCGGGTGTGCGTTGAGACCCAGAGCTCAGGACAGTGTACGAAGTCCCGCT 297
Db 181 GAAGACCTCGGGTGTGCGTTGAGACCCAGAGCTCAGGACAGTGTACGAAGTCCCGCT 240
QY 298 GTGGAAGTGGATGATCTGCTTCCATCACACTGCAAGTGTGCTGATGCTGCCCGGGAAC 357
Db 241 GTGGAAGTGGATGATCTGCTTCCATCACACTGCAAGTGTGCTGATGCTGCCCGGGAAC 300
QY 358 ATTTCTGCTCTCTGGGCTTTTGAATTTTGAATTTTGAATTTTGAATTTTGAATTTTGAATTT 417
```

Db 301 ATTTCTCTCTCGGCTCTTTAAGCAGCAGCTCCCTGAAATTGCCAGCCACATTTTGAATTA 360
Qy 418 CAAAAACAGAGGAGTTGTTTCCATGGTCAATTTTGAAATGACAGAAACCCCAAGCTGGAGAA 477
Db 361 CAAAACAGAGGAGTTGTTTCCATGGTCAATTTTGAAATGACAGAAACCCCAAGCTGGAGAA 420
Qy 478 TACCTACTTTTATTCAGAGTGAAGCTACCAATTAACAATATTTGTTTACAGTGAATATA 537
Db 421 TACCTACTTTTATTCAGAGTGAAGCTACCAATTAACAATATTTGTTTACAGTGAATATA 480
Qy 538 AGAATACCCCTCTTACACATTAAGAGCCTTACTTTAGAAATGAAACCAAGGAC 597
Db 481 AGAATACCCCTCTTACACATTAAGAGCCTTACTTTAGAAATGAAACCAAGGAC 540
Qy 598 GCCCTGGTCTGCATATCTCAGAGCGTTCCAGAGCCGATCGTGAATGGTGTCTTTGCCAT 657
Db 541 GCCCTGGTCTGCATATCTCAGAGCGTTCCAGAGCCGATCGTGAATGGTGTCTTTGCCAT 600
Qy 658 TCACAGGGGAAAGCTGTAAAGAAAGTCCAGCTGTTGTTAAAGAGGAGGAAAGTG 717
Db 601 TCACAGGGGAAAGCTGTAAAGAAAGTCCAGCTGTTGTTAAAGAGGAGGAAAGTG 660
Qy 718 CTTGATGAATTTTGGGAGGACATTAAGGTGCTGTCAGAAATGAATCGGCGAGGAA 777
Db 661 CTTGATGAATTTTGGGAGGACATTAAGGTGCTGTCAGAAATGAATCGGCGAGGAA 720
Qy 778 TGCACAGGCTGTTCAATAGATCTAAATCAAACTCCTCAGACCAATTTGCCACAATTA 837
Db 721 TGCACAGGCTGTTCAATAGATCTAAATCAAACTCCTCAGACCAATTTGCCACAATTA 780
Qy 838 TTTCTTAAAGTAGGGGAAACCTTTATGGATAAGGTGCAAGCTGTTTCATGTGAACCAATGA 897
Db 781 TTTCTTAAAGTAGGGGAAACCTTTATGGATAAGGTGCAAGCTGTTTCATGTGAACCAATGA 840
Qy 898 TTCGGGCTCACCTGGGAATTAGAANAACCAAGCACTCGAGAGGGCAACTACTTTGAGATG 957
Db 841 TTCGGGCTCACCTGGGAATTAGAANAACCAAGCACTCGAGAGGGCAACTACTTTGAGATG 900
Qy 958 AGTACTATTCAACAAACAGAACTATGATACGGATTCTGTTGCTTTGTTATCATCAGTG 1017
Db 901 AGTACTATTCAACAAACAGAACTATGATACGGATTCTGTTGCTTTGTTATCATCAGTG 960
Qy 1018 GCAAGAAACGACACCGGATATACACTGTTGTTCTTTCAAAGCATCCCAAGTCAATCAGCT 1077
Db 961 GCAAGAAACGACACCGGATATACACTGTTGTTCTTTCAAAGCATCCCAAGTCAATCAGCT 1020
Qy 1078 TTGTTTACCATCGTAGGAAAGGATTTAATATGCTTACCAATTCAGTGAAGTTATGAA 1137
Db 1021 TTGGTTACCATCGTAGGAAAGGATTTAATATGCTTACCAATTCAGTGAAGTTATGAA 1080
Qy 1138 ATTGACCAATATGAAGATTTTGTGTTTCTGTCAGGTTTAAAGCCTACCCACAATCAGA 1197
Db 1081 ATTGACCAATATGAAGATTTTGTGTTTCTGTCAGGTTTAAAGCCTACCCACAATCAGA 1140
Qy 1198 TGTAAGTGAACCTTCTTCGAAATCATTTTCTTTGTGAGCAAAAGGCTCTTGATAACGA 1257
Db 1141 TGTAAGTGAACCTTCTTCGAAATCATTTTCTTTGTGAGCAAAAGGCTCTTGATAACGA 1200
Qy 1258 TACAGCATATCCAGTTTGTCAATCATAGACCAAGCAGGAGGAGATATATATTCATGCA 1317
Db 1201 TACAGCATATCCAGTTTGTCAATCATAGACCAAGCAGGAGGAGATATATATTCATGCA 1260
Qy 1318 GAAATGATGATGCCCAATTTTACCAAAATGTTTCACGCTGAATATAAGAGGAACCTCAA 1377
Db 1261 GAAATGATGATGCCCAATTTTACCAAAATGTTTCACGCTGAATATAAGAGGAACCTCAA 1320
Qy 1378 GTGCTCGCAGAGACATCGGCAAGTCAAGGCTGCTGTTTCTCGGATGGATACCAATACCA 1437
Db 1321 GTGCTCGCAGAGACATCGGCAAGTCAAGGCTGCTGTTTCTCGGATGGATACCAATACCA 1380
Qy 1438 TCTTGACCTGGAAGAGTGTTCAGACAAGTCTCCCAACTGCAACAGAGATCAAGAA 1497
Db 1381 TCTTGACCTGGAAGAGTGTTCAGACAAGTCTCCCAACTGCAACAGAGATCAAGAA 1440

Qy 1498 GGAGTCTGGAATFAGAAAGCTTAACAGAAAGTGTGTTGGAAGTGGGTGTGAGCAGTACT 1557
Db 1441 GGAGTCTGGAATFAGAAAGCTTAACAGAAAGTGTGTTGGAAGTGGGTGTGAGCAGTACT 1500
Qy 1558 CTAACATGAGTGAAGCCATAAAGGGTTCCTGGTCAAGTGTGTCATACAAATCCCTT 1617
Db 1501 CTAACATGAGTGAAGCCATAAAGGGTTCCTGGTCAAGTGTGTCATACAAATCCCTT 1560
Qy 1618 GGCACATCTTGTGAGACGATCCTTTTAAACTCTCCAGGCCCTTCCCTTTTCAATCCAGAC 1677
Db 1561 GGCACATCTTGTGAGACGATCCTTTTAAACTCTCCAGGCCCTTCCCTTTTCAATCCAGAC 1620
Qy 1678 AACATCTCAATCTATGCAACAAATTTGGTGTGTTCTCTCTTTCATTTGCTTTTAAACCTG 1737
Db 1621 AACATCTCAATCTATGCAACAAATTTGGTGTGTTCTCTCTTTCATTTGCTTTTAAACCTG 1680
Qy 1738 CTAATTTGTCAAGTACAAAAGCAATTTAGGTATGAAGCCAGCTACAGATGGTACAG 1797
Db 1681 CTAATTTGTCAAGTACAAAAGCAATTTAGGTATGAAGCCAGCTACAGATGGTACAG 1740
Qy 1798 GTGACCGGCTCTCAGATAATCAGTACTTCTACGTTTGTATTTTCAAGAGAATATGAATATGAT 1857
Db 1741 GTGACCGGCTCTCAGATAATCAGTACTTCTACGTTTGTATTTTCAAGAGAATATGAATATGAT 1800
Qy 1858 CTCAATGGGAGTTCCTCAAGAGAAATTTTAGAGTTTGGGAAGTACTTAGGATCAGGTGCT 1917
Db 1801 CTCAATGGGAGTTCCTCAAGAGAAATTTTAGAGTTTGGGAAGTACTTAGGATCAGGTGCT 1860
Qy 1918 TTTGAAAGATGATGAACGCAACAGCTTATGGAAATTTAGCAAAACAGGAGTCTCAATCCAG 1977
Db 1861 TTTGAAAGATGATGAACGCAACAGCTTATGGAAATTTAGCAAAACAGGAGTCTCAATCCAG 1920
Qy 1978 GTTGGCGTCAAAATGCTGAAGAAAGAGCAGACAGCTCTGAAGAGAGAGGACACTCATGTCA 2037
Db 1921 GTTGGCGTCAAAATGCTGAAGAAAGAGCAGACAGCTCTGAAGAGAGAGGACACTCATGTCA 1980
Qy 2038 GAACTCAAGATGATGACCCAGCTGGGAAGCCAGAGAAATTTGTGAACCTGCTGGGGGG 2097
Db 1981 GAACTCAAGATGATGACCCAGCTGGGAAGCCAGAGAAATTTGTGAACCTGCTGGGGGG 2040
Qy 2098 TGCACTGTGAGGACCAATTTTACTGTTTGAATATCTGTTGCTATGGTGAATCTTCTC 2157
Db 2041 TGCACTGTGAGGACCAATTTTACTGTTTGAATATCTGTTGCTATGGTGAATCTTCTC 2100
Qy 2158 AACTATCTAAGAGTAAAGAGAAATTTTCAAGACCTTGAGACAGAGATTTTCAAGGAA 2217
Db 2101 AACTATCTAAGAGTAAAGAGAAATTTTCAAGACCTTGAGACAGAGATTTTCAAGGAA 2160
Qy 2218 CACAATTTTCAAGTTCCTTCAATCACAATCACAATTTCCAGCATGCTGGTTCA 2277
Db 2161 CACAATTTTCAAGTTCCTTCAATCACAATCACAATTTCCAGCATGCTGGTTCA 2220
Qy 2278 AGAAGATTCAGATACACCCGACCTCGGATCAAAATCTCAGGCTTCATGGGAATTCATTT 2337
Db 2221 AGAAGATTCAGATACACCCGACCTCGGATCAAAATCTCAGGCTTCATGGGAATTCATTT 2280
Qy 2338 CACTCTGAAGATGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2397
Db 2281 CACTCTGAAGATGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2340
Qy 2398 AATGTCCTTACATTTGAAGATCTTTTCTGTTTGAATATCAAGTTCGCAAGGAATGAA 2457
Db 2341 AATGTCCTTACATTTGAAGATCTTTTCTGTTTGAATATCAAGTTCGCAAGGAATGAA 2400
Qy 2458 TTTCTGGAATTTAAGTGTGTGTTCAGAGACCTGGCCGCCAGGAAACGTGTTGTCAAC 2517
Db 2401 TTTCTGGAATTTAAGTGTGTGTTCAGAGACCTGGCCGCCAGGAAACGTGTTGTCAAC 2460
Qy 2518 CACGGGAAAGTGGTGAAGATATGTGATTTGATTTGGCTCGAGATATCATGAGTATTC 2577
Db 2461 CACGGGAAAGTGGTGAAGATATGTGATTTGATTTGGCTCGAGATATCATGAGTATTC 2520

1361 AGTTTACCAAAATGTTTACACGCTGAATATAAGAAAGAAACCTCAAGTCTAGCAAAATGCCT 1420 Db
1394 CGGCAAGTCAGGCGCTCTGTTTCTCGATGGATACCCATTACCATCTTGGACCTTGAAGA 1453 Qy
1421 CAGCCAGCAGGCGCTCTGTTTCTCTGATGGCTACCCGCTACCTCTTGGACCTTGAAGA 1480 Db
1454 AGTGTTCAGACAAGTCTCCCAACTGCAACAAGAGATCACAGAAGGAGTCTGGAATAGAA 1513 Qy
1481 AGTGTTCGACAAATCTCCCAATGCAACGAGAAATCCCAAGAGGAGTCTGGAATAAAA 1540 Db
1514 AGGCTAACAGAAAGTGTTCGACAGTGGGTGTCAGCAGTACTCTTAAACATGAGTGAAG 1573 Qy
1541 AGGCTAACAGAAAGTGTTCGACAGTGGGTGTCAGCAGTACTCTTAAATATGAGTGAAG 1600 Db
1574 CCATAAAGGGTCTCTGCTCAAGTGTGTCATACAAATCCCTTGGCACATCTTGTGAGA 1633 Qy
1601 CCGGAAAGGGCTCTCTGCTCAAAATGCTGTCGCTACAAATCTATGGGCACGCTCTTGGAAA 1660 Db
1634 CGATCCTTTTAAACTCTCCAGGCGCCCTTCCCTTTTCATCCAGACAACATCTCAATCTATG 1693 Qy
1661 CCATCTTTTAAACTCACAGGCGCCCTTCCCTTTTCATCCAGACAACATCTCTCTTCTATG 1720 Db
1694 CAACAATGTGTGTCTCTCTCTTCAATGTGTGTTTAAACCTGCTTAATTTGTCAACAAGT 1753 Qy
1721 CGACCAATGGGCTCTGTCTCCCTTCAATGTGTGTTTCTCATTTGTGTGATCTGCCACAAAT 1780 Db
1754 ACBAAGCAATTTAGGTATGAAGCAGCTACAGATGTTACAGGTGACCGGCTCTCTCAG 1813 Qy
1781 ACAAAGCAATTTAGGTATGAAGCAGCTACAGATGTTACAGGTGACCTGGCCCTCTGG 1840 Db
1814 ATAATGAGTACTTCTACGTTGATTTACAGAAATATGAATATGATCTCAAAATGGAGTTTC 1873 Qy
1841 ATAAAGGACTTCTACGTTGATTTACAGGACTATGAATATGATCTTAAAGTGGAGTTTC 1900 Db
1874 CAAGAGAAATTTAGGTATGGAAAGGTA CTAGATCAGGTGCTTTTGGAAAAAGTGATGA 1933 Qy
1901 CGAGAGAACTTTAGGTATGGAAAGGTCCTGGGCTCTGGGCTCTGGGAGGTTGATGA 1960 Db
1934 ACCCAAGCTTATGGAATTTAGCAAAACAGAGTCTCAATCCAGTTGCGCTCAAAATGC 1993 Qy
1961 ACCCCAGGCTATGGCAATTTAGTAAACCGGAGTCTCAATTTAGGTGGCGGTGAAGATGC 2020 Db
1994 TGAAGAAAAAGCAGACAGCTCTGAAAGAGAGGCACTCATGTGAGAACTCAAGATGATGA 2053 Qy
2021 TAAAGAAAGGCTGACAGCTGTGAAAGAAAGAGCTCTCATGTGCGAGCTCAAAATGATGA 2080 Db
2054 CCAGCTGGGAAGCCACGAGAAATTTGTGAACTGTGCTGGGGGCTGCACACTGTGAGGAC 2113 Qy
2081 CCCACCTGGGACCACTGACAACTGTGAATCTGTGGGGCATGACACTGTGAGGGC 2140 Db
2114 CAATTTACTTGAATTTGAATCTGTTGCTATGTTGATCTTCAACTATCTAAGAGTA 2173 Qy
2141 CAGTGTACTTGAATTTTGAATATTTGTTGCTATGTTGACCTCTCACTACTAAGAGTA 2200 Db
2174 AAAGAGAAAAATTTACAGGACTTGGACAGAGATTTTCAAGGAACAAATTTTCAAGTTT 2233 Qy
2201 AAAGAGAAAGTTTACAGGACATGGACAGAGATTTTAAAGAACATAAATTTTCAAGTTT 2260 Db
2234 ACCCCACTTTCCAAATCACTCCAAATCCAGCATGCTGTTTCAAGAGAAAGTTTCAAGATC 2293 Qy
2261 ACCCTACTTTCCAGGACACTTCAAAATCCAGCATGCTGTTTCAAGAGAAAGTTTCAAGTTT 2320 Db
2294 ACCCGNCTCGGATCAAACTCAGGCTTCAATGGGAATTCATTTCACTCTGAAGATGAAA 2353 Qy
2321 ACCCGCTCTGGATCAGCTCTCAGGCTTCAATGGGAATTCATTTCACTCTGAAGATGAGA 2380 Db
2354 TTGAATATGAAACCAAAAAAGGCTG-----GAAGAGAGGAGGACTTGAATGTGCTTA 2407 Qy
2381 TTGAATATGAAACCAAGAGGCTGGCAGAAAGAGGAGGAGATTTGAACTGTCTGA 2440 Db
2408 CATTTGAAGATCTCTTTTGTCTTGTGCTATCATCAAGTTGCAAGGAAATGGAATTTCTGGAAT 2467 Qy

2441 CGTTTGAGACCTCTCTTTGCTTTGCGTACCAAGTGGCCAAAGCATGGAATTTCTGAGT 2500 Db
2468 TTAAGTCGTGTGTTTACAGAGACCTGGCCGCCAGGAACGTGTTGTCAACCCACGGGAAG 2527 Qy
2501 TCAAGTCGTGTGTTTACAGAGACCTGGCAGCAGGAATGTGTTGTCAACCCACGGGAAG 2560 Db
2528 TGTGTAAGATATGTGATTTTGGATTTGGCTCGAGATATCATGAGTGATTCAACTATGTTG 2587 Qy
2561 TGTGTAAGATCTGTGATTTTGGATTTGGCTCGCAGACATCTCTGAGCGACTCCAGCTACGTCG 2620 Db
2588 TCAGGGGCAATGCCCTCTGCTGTAATGGAATGGCCCGGCAAGGCTGTGTTCAAGGCA 2647 Qy
2621 TCAGGGGCAACGCAACGCTGCGGTGAGTGGATGGCACCCGAGAGCTTATTTGAAGGGA 2680 Db
2648 TCTACACCAATTAAGAGTGATGCTGTGTCATATGGAATATTAATCTGTGGAAATCTTCTCAC 2707 Qy
2681 TCTACCAATCAAGAGTGAGCTCTGTGCTCTACGGCATCTCTCTGCGAGATATTTTTCAC 2740 Db
2708 TTGGTGTGAATCTTACCTGGCATTTCCGTTGATGCTAACTTCTCAAACTGATTTCAA 2767 Qy
2741 TGGGTGTGAACCTTACCTGGCATTTCTGTGACGCTAACTTCTATATAAACTGATTCAGA 2800 Db
2768 ATGGATTTAAATGGAATCAGCCATTTTATGCTACAGAAAGAAATATACATATATATGCAAT 2827 Qy
2801 GTGGATTTAAATGGAAGCGCCATTTCTATGCCACAGAAAGGATATATCTTTGTAATGCAAT 2860 Db
2828 CTGTCTGGGCTTTGACTCAAGGAAACGGCCATCTTCCCTAAATTTTGAATTTTGTAG 2887 Qy
2861 CTGTCTGGGCTTTGACTCAAGGAAAGCGCCATCTTCCCAACCTGACTTCAATTTTGTAG 2920 Db
2888 GATGTGAGCTGGCAGATGCAAGAAAGCGATGATATGAAATGTGGATGGCGGTGTTGCG 2947 Qy
2921 GATGTGAGCTGGCAGAGGCAAGAAAGCGATGATATGAAACATGGGTGGCAACGCTCCAG 2980 Db
2948 AATGTCTCACACTACCAACCAAGCGACTTTTACAGAGAGATGATTTGGGGCTAC 3007 Qy
2981 AACATCACTCACTTACCAAAACAGGCGCCCTCTACAGAGAGGCGGCTCAGAGCCGC 3040 Db
3008 TCTCTCGGAGGCTCAGGTGAAAGATTCGTAGAGGAACAAATTTAGTTTAAAGACTTTCAT 3067 Qy
3041 CATCGCCACAGGCGCCAGGT- GAAGATTCACAGAGAAAGATTTAGCAGGAGGCTTTGGA 3099 Db
3068 CCTTCCACCTATCTTAAACAGGCTGTAGATTA CAAAGAAAGATTAATTTATCATCACTAAA 3127 Qy
3100 CCCCAGCA-----CCCTAGCAGGCTGTAGACCGCAGAGCCAAGATTAGCTCTCGCTCT- -G 3153 Db
3128 AGAAATCTATATCACTGCTCTTCAACAGACTTTTCTCTAGAGCGCTCTGCGTTTA 3187 Qy
3154 AGGAAGCGCCCTACAGGCGGTGCTTCTGCTGGAATTTTCTCTAGATGCTGTCTGCCATTA 3213 Db
3188 CTCTTGTGTTTCAAGGAGCTTTTGTAAATCAAAATCAATCATCTGTCAAGGCGAGGAGGC 3247 Qy
3214 CTC-----CAAAGTGACTTCTATAAAATCAAACTCTCTCGCAGCGGCGAGAGCC 3266 Db
3248 TGATAATGAATTTATTTGAGCAATTTGATCTGCATCCAGGCTTCTCAGGCGGCTTTGAG 3307 Qy
3267 AATAATGAGACTTTTGGTGAGCGCCCTACCTCGGGGCGCTTTTCCAGGCGCCCGCAGC 3326 Db
3308 TGAATTTG-----TGTACTGAGTACAGTATATTTCTTGTAAATACATAAAACAAA-- 3357 Qy
3327 TTGAGGGGAAGCATGTATCTGAAATATAGTATATTTCTTGTAAATACGTGAACCAACC 3386 Db
3358 -----AGCATTTTGTAAAGGAGAGCTAATATGATTTT-----AAGTCTATGTTTAAAT 3410 Qy
3387 AAACCCGTTTTTGTCTAAGGAAAGCTTAAATATGATTTTAAATATCTATGTTTTTAAAT 3446 Db
3411 AATATGTAAATTTTTCAGCTTATTTAGTGATATATTTTATGGTGGGAATATAAATTTCTAC 3470 Qy
3447 ACTATGTACTTTTTCATCTATTTAGTGATATATTTTATGATGATAATTTATGATGGAATAAATTTCTAC 3506 Db
3471 TACAGAAAAA 3485 Qy
3507 TGTAGAAAAA 3521 Db

RESULT 14	MMFLT3	3521 bp	mRNA	linear	ROD 04-DEC-1991
LOCUS	Mouse Flt3 mRNA for tyrosine kinase receptor of the PDGF.				
DEFINITION	X59398				
ACCESSION	X59398.1	GI:50978			
VERSION	Flt3 gene; tyrosine kinase receptor.				
KEYWORDS	Mus musculus (house mouse)				
SOURCE	Mus musculus				
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
REFERENCE	1	Rosnet, O., Marchetto, S., delapeyriere, O. and Birnbaum, D.			
AUTHORS	Murine Flt3, a gene encoding a novel tyrosine kinase receptor of				
TITLE	the PDGFR/CSF1R family				
JOURNAL	Oncogene 6 (9), 1641-1650 (1991)				
MEDLINE	92019834				
PUBMED	1656368				
REFERENCE	2 (bases 1 to 1321)				
AUTHORS	Rosnet, O.				
TITLE	Direct Submission				
JOURNAL	Submitted (07-MAY-1991) O. Rosnet, Cancerologie et therapeutique exp., Inst National de la Sante, 27 Boulevard Lei Roure, 13009 Marseille, France				
FEATURES	Location/Qualifiers				
source	1..3521				
	/organism="Mus musculus"				
gene	/mol_type="mRNA"				
mRNA	/db_xref="taxon:10090"				
	1..3521				
	/gene="Flt3"				
	1..3521				
	/gene="Flt3"				
CDS	/evidence="experimental"				
	82..3084				
	/gene="Flt3"				
	/codon_start=1				
	/product="Flt3 protein"				
	/protein_id="CAA42041.1"				
	/db_xref="GI:50979"				
	/db_xref="MGI:95559"				
	/db_xref="SWISS-PROT:Q00342"				
	/translation="MRALAQRDRRLLLVLSVMILETVNQDLPIVKCLVLSHENV GSSAGKPSYRMVRGSPEDLQCTPRQSEGVTEAATVEVASEGSITLQVLATPGDL SCLWFKHSLSCQPHFDLQNGRGIYSMALLVNTEQAGEYLLHIQSEANYYTLPTVN VPTDLVYLRRPYFRKMENQDALLCISGEVPEPTVWVLCSSHRESKEEGPAVRKE EKVLHFGTDIRCCARNALGRESHTLFDLNQAPSTLPQLFKVGEPLWTRKAI KHPGSLATWLEEDKALGESYFEMSTYNTMRIRILLAFVSVGRNDTYGYTSSS YHNSQALYTLLEKGIINATSSQEYEDPYKEFCFVRKAYPRIRCTWIFSOASP CBQGLDEGYSYKFCSDHKNKPEYIFVAENDAOFTKMTLNIRKPKOVLNANASQ ASCSDGYPLPSWTWKCKSDSPNCTEIPSGVWNKKNRKFQGWVSSSTLANSEAG KGLVKCCAYNSMGTSCETIFLNSPGPPFTQDNISFYATIGLCLPIVLVLICHK YKQFRYESQLQMTQVTGLDNEYFYVDFRDEYDLKWEFFPRENLKFGVLGSAFOR VNNATYKSGTQISQVAVMLKEKADSCKEALMSLKMTHLGHNDNI VNLGAC TSLGVPYLIFFCYCGDLLNLYRSKREPHRTWTEI FKEHNFSPYPTFAHSSNSMPG SREVQLHPLDQLSGFNQNSIHSEDEI EYENOKRLAEBEEDNLVLFEDLLCPAYOV AKGMFLPEKSCVHRDLAARNVLVTHGVKVKICDQPLGARDILSDSSVYVRGNALPVK WNAPESLFGIYTIKSDVWSYGILLWEIPLSGVNPYFGIPVDANFYKLIQSGFKNEQP FYATGEGIVFMQSCWAFDSRKRPSPFNLTSLGCOLAEAEAEAMYNQMGNNVPHEPSIY QNRRLPSRAGSEPPSPQAVKIHRSR"				
BASE COUNT	943 a	854 c	907 g	817 t	
ORIGIN					
Query Match	65.0%;	Score 2274.6;	DB 10;	Length 3521;	
Best Local Similarity	80.3%;	Pred. No. 0;			
Matches 2805;	Conservative 0;	Mismatches 649;	Indels 41;	Gaps 10;	
QY	17	GGCTGGCGCGCGCCCTGGGGGACCCCGGGCTCCGGAGGCCATCCGCGCGTTGGC---GC	73		
Db	42	GGCTGGCGCGCGGCT-GGCTACCGCGCTCCGGAGGCCATCGCGCGTTGGCGGCG	100		

1181 AGTTCTGCTCTCAGTCAGGTTTAAAGCGTACCCAGAAATCCGATGCACTGATCTTCT 1240
1214 CTCGAAATCATTTCTTGTGAGCAAAAGGCTCTTGATACCGATACAGCATATCCAAGT 1273
1241 CTCAGGCTCATTTCTTGTGAAACAGAGGCTTGGAGGATGGGTACAGCATATCTAAAT 1300
1274 TTGCAATCATAGCACCGACCGAGGAATATATATTTCCATGAGAAAATGATGATGCC 1333
1301 TTGCGATCATAGAACAGCCAGGAGGTACATATTTCTATGACAGAAAATGATGATGCC 1360
1334 AATTTACCAAAATGTTCAAGCTGAATATAGAGGAACCTCAAGTCTCGAGAGCAT 1393
1361 AGTTCAACCAAAATGTTCAAGCTGAATATAGAGGAACCTCAAGTCTCGAGAAATGCT 1420
1394 CGCAAGTCAGGCTGCTGTTCTCGATGATACCATTTACCATTCTTGACCTGGAAGA 1453
1421 CAGCCAGCCAGGCTGCTGTTCTCTGATGCTTACCCGCTACCCCTCTTGACCTGGAGA 1480
1454 AGTGTTCAGACAAAGTCTCCCAACTGCAAGAGAGATCACAGAGGAGTCTGGAATAGAA 1513
1481 AGTGTTCGGACAAATCTCCCAATTGCAACGGAGGAATCCAGAGGAGTCTGGAATAAAA 1540
1514 AGCTTAACAGAAAAGTGTGTTGACAGTGGGTGTCGAGCAGTACTCTAAACATGAGTGAAG 1573
1541 AGGCTAACAGAAAAGTGTGTTGCGCAGTGGGTGTCGAGCAGTACTCTAAATATGAGTGAAG 1600
1574 CCATAAAAGGTTCTGTTCAAGTGTGTCATACAAATCCCTTGGCACATCTTTGTGAGA 1633
1601 CCGGAAAGGCTTCTGTTCAATGCTGTCGATCAAAATCTATGGGACGCTCTTGGAAA 1660
1634 CGATCCTTTTAACTCTCCAGGCCCCCTTCCCTTTTCATCCAGACAAACATCTCAATTCATG 1693
1661 CCATCTTTTAACTCACAGGCCCTTCCCTTTTCATCCAGACAAACATCTCTCTTCTATG 1720
1694 CAACAAATGTTGTTCT 1753
1721 CGACCAATGGGCTCTGTCTCCCTCTCAATGTTGTTCTCTCTCTCTCTCTCTCTCTCTCT 1780
1754 ACAAAGCAATTTAGTATGAAGCCAGCTACAGATGGTACAGGTGACCGGCTCTCTCAG 1813
1781 ACAAAGCAATTTAGTATGAAGCCAGCTACAGATGGTACAGGTGACCGGCTCTCTCAG 1840
1814 ATATGAGTACTTCTAGTTGATTTTCAGAGAATATGAATATGATCTCAAAATGGAGTTTC 1873
1841 ATAAAGGACTTCTAGTTGATTTTCAGGAGCTATGAATATGATCTTAAAGTGGAGTTCC 1900
1874 CAAGAGAAAATTTAGTATGGGAGGTACTAGGATCAGGTGCTTTTGGAAAAGTGTATGA 1933
1901 CGAGAGAGAACTTAGAGTTTGGGAAAGTCTCTGGGCTCTGGGCTTTCGGGAGGTTGATGA 1960
1934 ACCCAAGCTTATGGAATTTAGCAAAACAGAGCTCTCAATCCAGGTTGGCTCAAAATGC 1993
1961 ACCCAAGCTTATGGAATTTAGCAAAACAGAGCTCTCAATCCAGGTTGGCTCAAAATGC 2020
1994 TGAAGAAAAGCAGACAGCTCTGAAAAGAGAGGCACTCATGTACAGAACTCAAGATGATGA 2053
2021 TAAAGAAAAGCTGACAGCTGTGAAAAGAGAGCTCTCATGTCCGAGCTCAAAATGATGA 2080
2054 CCAGCTGGGAAGCCAGAGAAATTTGTGAACCTGTCTGGGGGCGTGACACTGTGAGGAC 2113
2081 CCCACCTGGGACACCATGACACATCGTGAATCTGTCTGGGGGCGTGACACTGTGAGGCG 2140
2114 CAATTTACTTTGATTTTGAATCTGTTGCTATGTTGATCTCTCAACTATCTAAGAGTA 2173
2141 CAGTGTACTTTGATTTTGAATATTTGTTGCTATGTTGATCTCTCAACTATCTAAGAGTA 2200
2174 AAAGAGAAAATTTACAGCACTTGGACAGAGATTTTCAAGGAACACAAATTTCAAGTTTTT 2233
2201 AAAGAGAGATTTTACAGACATGAGACAGATTTTAAAGAACATAAATTTCAAGTTTTT 2260
2234 ACCCCACTTTTCCAATCATCTCCAAATTTCCAGCATGCTGTTTCAAGAGAGTTTCAGATAC 2293

2261 ACCCTACTTTCCAGGCACATTTCAAAATTCAGCATGCTGTTCCAGAGAGTTCAAGTTAC 2320
2294 ACCCGACTCGGATCAAAATCTCAGGGCTTTATGGGAATTCATTTCACTCTGAAGATGAAA 2353
2321 ACCCGCTTTGGATCAGCTCTCAGGGTTCAATGGGAATTCATTTCACTCTGAAGATGAGA 2380
2354 TTGAATATGAATAACCAAAAAGGCTG-----GAAGAAGAGGAGGACTTGAATGTGCTTA 2407
2381 TTGAATATGAATAACCAAAAAGGCTTGGCAGAAAGAGAGGAGGAGATTTGAACGTGCTGA 2440
2408 CATTTCAAGATCTTTCTTTGCTTTGCTATCAAGTGTCCAAAGGAATGGAATTTCTGCAAT 2467
2441 CGTTTGAAGACCTTCTTTGCTTTGCTATCCAGTGGCCAAAGCATGGAATTTCTGAGT 2500
2468 TTAAGTCGTGTGTTCACAGAGACTTGGCCGCGCAGAGAACTGCTTGTCAACCACCGGAAAG 2527
2501 TCAAGTCGTGTGTTCACAGAGACTTGGCCGCGCAGAGAACTGCTTGTCAACCACCGGAAAG 2560
2528 TGGTGAAGATATGTGACTTTTGGATTTGGCTCGAGATATCATGAGTATTCAACTATGTTG 2587
2561 TGGTGAAGATCTGTGACTTTTGGACTGGCCGCGCAGACATCTCTGAGCGACTCCAGCTACGTCG 2620
2588 TCAGGGGCAATGCTGCTGCTGTAATAATGGATGGCCCGCCGAAAGCTCTGTTGAAGGCA 2647
2621 TCAGGGGCAACCGCACCGGCTGCGGTGAAGTGGATGGCACCCGAGAGCTTATTTGAAGGGA 2680
2648 TCTACACCAATTAAGAGTGTCTGCTCATATGGAATATTTACTGTGGGAAATCTTCTCAC 2707
2681 TCTACACATCAAGAGTGTGCTGCTCTACGGCATCTTCTCTGGGAGATATTTTCTCAC 2740
2708 TTGGTGTGAATCTCTTACCTTGGCATTTCCGGTTGATGCTAACTCTCTACAACTGATTCAAA 2767
2741 TGGGTGTGAACCTTACCTTGGCATTTCTGTGCGAGCTAACTTCTATAAACTGATTCAGA 2800
2768 ATGGATTTAAATGGATCAGCCATTTTGTCTACAGAGAAATATACATTAATGCAAT 2827
2801 GTGGATTTAAATGGAGCGCCATTTCTATGCCACAGAGGATATATCTTTGTAATGCAAT 2860
2828 CTGCTGGGCTTTTGACTCAAGGAAACGGCCATCTTCCCTAAATTTGACTTTCGTTTTTAG 2887
2861 CTGCTGGGCTTTTGACTCAAGGAAAGGGCCATCTTCCCAACCTGACTTTCATTTTTTAG 2920
2888 GATGTCAGCTGGCAGATGACAGAAAGCGATGTATCAGAAATGTGGATGGCGGTGTTTTCGG 2947
2921 GATGTCAGCTGGCAGAGGAGAGAGCGATGTATCAGAACTGGGTGGCAACGCTCCAG 2980
2948 AATGTCCTCACACTTACCAAAACAGGCGACTTTTTCAGCAGAGATGGATTTGGGGCTAC 3007
2981 AACATCCATCTTCTACCAAAACAGGCGCCCTCTCAGCAGAGAGGGGCTCAGAGCGC 3040
3008 TCTCTCCGAGGCTCAGGTGCAAGATTTCTGAGAGGAACAATTTTAGTTTAAAGGACTTTCAT 3067
3041 CATGGCAACAGGCCCAAGT-GAAGATTCACAGAGAAAGAGTTAGCGAGAGGCTTTGGA 3099
3068 CCCTCCACCTTATCCCTAACAGGCTGTAGATTTACCAAAACAGAGATTAATTTTCACTCAAAA 3127
3100 CCCCGCCA-----CCCTAGCAGGCTGTAGACCGCAGAGCAAGATTAGCCTCGCCTCT-G 3153
3128 AGAAATCTATATATCACTGTCTCTCAACAGATTTTCTCTAGAGCCGCTGTGCGTTTA 3187
3154 AGGAAGCGCCCTACAGGCGGTTGCTTTCGCTGGAGCTTTTCTCTAGATGCTGTCTGCTCA 3213
3188 CTCTGTTTTTCAAGGAGCTTTTGTAAATCAATCATCTCTGTCACAGAGCGAGGAGC 3247
3214 CTC-----CAAGTGTACTTCTATAAATCAAACTCTCTCTGCAACAGCGGGAGAGCC 3266
3248 TGATAATGAATTTTATTTGGAGCATTTGATCTGCACTCAAGGCTTCTCAGGCGGCTTTGAG 3307
3267 AATAATGAGACTTGTGTTGGAGCCCGCTTACCTTGGGGGCTTCTCCAGGCCCCCAGGC 3326
3308 TGAATTTG-----TGATCCTGAAGTACAGTATATTTCTGTAAATACATAAAACAAA-- 3357
3327 TTGAGGGGAAAGCCATGTATCTGAAATATAGTATATTTCTTGTAAATACGTGAAACAAACC 3386

QY	3358	-----AGCATTTTGTCTAAGGAGAGCTAATANGATTTTTT--AAAGTCTATGTTTTTAAAAAT	3411
Db	3387	AAACCCGTTTTTGTCTAAGGAGAGCTAAATATGATTTTTTAAAAATCTATGTTTTTAAAAAT	3446
QY	3411	AATATGTAATTTTTTTCAGCTATTATAGTGTATATTTTTATGGTGGGAATAAAATTTCTAC	3470
Db	3447	ACTATGTAACCTTTTTCATCTATTTATGATATATTTTTATGATGGAATAAACTTTCTAC	3506
QY	3471	TACAGAAAAAATAA 3485	
Db	3507	TGTAGAAAAAATAA 3521	
RESULT 15			
AR005211	AR005211	Sequence 1 from patent US 5747651.	
DEFINITION	AR005211	Sequence 1 from patent US 5747651.	
ACCESSION	AR005211		
VERSION	AR005211.1	GI:3966090	
KEYWORDS	Unknown.		
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	Unclassified.		
AUTHORS	1 (bases 1 to 3453)		
TITLE	Lemischka, I.R.		
JOURNAL	Antibodies against tyrosine kinase receptor flk-1		
FEATURES	Patent: US 5747651-A 1 05-MAY-1998;		
source	Location/Qualifiers		
	1. 3453		
	/organism="unknown"		
BASE COUNT	947 a 822 c 875 g	809 t	
ORIGIN			
Query Match 63.6%; Score 2227.2; DB 6; Length 3453;			
Best Local Similarity 79.6%; Pred. No. 0;			
Matches 2775; Conservative 0; Mismatches 653; Indels 56; Gaps 10;			
QY	35	GGGAGCCCCGGCTCCGAGGCCATCCGGCGTTGGC----	91
Db	8	GGCTACCGCGCTCCGAGGCCATCGGGCGTTGGCGAGCGACCGCGCGCTGC	67
QY	92	CGCTGCTCGTTGTTTTTCTGCAATGATATTTGGGACTATTACAATCAAGATCTGCCTG	151
Db	68	TGCTGCTGTTGTTTGTTCAGTAATGATCTTGAGACCGTTACAAACCAAGACCTGCCTG	127
QY	152	TGATCAAGTGTGTTTTTAATCAATCATAAGAACAAATGATTCATCAGTGGGAAAGTCATCAT	211
Db	128	TGATCAAGTGTGTTTTTAATCAGTCATAGAACAAATGGCTCATCAGCGGAAAGCCATCAT	187
QY	212	CATATCCCATTGTTATCAGAAATCCCGAGAGCTCGGGTGTGGTTGAGACCCAGAGCT	271
Db	188	CGTACCGAATGTGCGAGGATCCCGAGAGACTCCAGTGTATACCCGAGGCGCCAGAGTG	247
QY	272	CAGGACAGTGTACGAAGCTCGCGTGTGGAAAGTGGATGTATCTGCTTCCATCACACTGC	331
Db	248	AAGGACCGGTATATGAAGCGGCCACCGTGTGGTCCAGTCTGGTTCATCACCTGC	307
QY	332	AAAGTGTGTCGATGCCCGAGGACATTTCTGCTCTCTGGGTCTTTAAGCACAGCTCCC	391
Db	308	AAAGTGTGTCGATGCCCGAGGACATTTCTGCTCTCTGGGTCTTTAAGCACAGCTCCC	367
QY	392	TGAATTTGCCAGCCACATTTTGTATTTACAAACAGAGGAGTGTCTTTCCATGTCATTGTA	451
Db	368	TGGGCTGCCAGCGGACATTGATTTACAAACAGAGGAAATCGTTTCCATGGCCATCTTGA	427
QY	452	AAATGACAGAAACCAAGCTGGAGAAATACCTACTTTTTTATTCAGAGTGAAGCTACCAATT	511
Db	428	ACGTGACAGAGACCCAGGACGAGAAATACCTACTCCAATTCAGAGCGAAACGCGCCAACT	487
QY	512	ACACAATATTTTACAGTGTAGTATAAGAAATACCTCTGTTTACATTAAGAGACCTT	571
Db	488	ACACAGTACTGTTTACAGTGTATGAAGATATACAGCTGTATGTGTGAAGAGACCTT	547

1652 QY CAGGCCCTTCCCTTTCATCCAGAGCAACATCTCATCTTATGCAACAATTTGGTGTGTC 1711
 1628 Db CAGGCCCTTCCCTTTCATCCAGAGCAACATCTCTTCTATGCGACCATTTGGGCTCTGTC 1687
 1712 QY TCCTCTTCACTTCGCTTTTAAACCTGCTAATTTGTACAGAGTACAAAAGCAATTTAGGT 1771
 1688 Db TCCTCTTCACTTCGCTTTTCAATTTGTTGATCTGCCACAAATACAAAAGCAATTTAGGT 1747
 1772 QY ATGAAGCCAGCTACAGATGGTACAGGTGACCGGCTCTCAGATATGAGTACTTCTACG 1831
 1748 Db ACAGAGTCACTGTCAGATGATCCAGTGACTGGCCCTCGATACAGTACTTCTACG 1807
 1832 QY TTGATTTACAGAAATATGAATATGATCTCAAAATGGGAGTTTCCACAGAGAAATTTAGAGT 1891
 1808 Db TTGACTTCAGGAGTATGAATATGATCTTAAAGTGGGAGTTCCCGAGAGAACTTAGAGT 1867
 1892 QY TTGGGAGGACTAGGATCAGGTGCTTTTGGAAAGTGTGATGACGCAACAGCTTATGGAA 1951
 1868 Db TTGGGAGGACTTGGGCTCTGGGCTTTCCGGAGGAGTGAACCGCACGGCCCTATGGCA 1927
 1952 QY TTAGCAAAAACAGAGTCTCAATCCAGGTTGCCCTCAAAATGCTGAAAGAAAAGCAGACA 2011
 1928 Db TTAGTAAACCGGAGTCTCAATTCAGTGGGCTGAAGTCTTAAAGAGAGCTTGACA 1987
 2012 QY GCTCTGAAAGAGAGGCACTCATGTCAAACTCAAGATGATGACCCAGCTGGGAAGCCAG 2071
 1988 Db GCTGTGAAAGAGAGCTCTCATGTGGAGCTCAAAATGATGACCCAGCTGGGACACCATG 2047
 2072 QY AGAATATTGTGAACCTGCTGGGGCTGCACACTGTCAGGACCAATTTACTGATTTTG 2131
 2048 Db ACAACATCTGAAATCTGCTGGGGCTGCACACTGTCAGGCGCCAGTGTACTGATTTTG 2107
 2132 QY AATACTGTTGCTATGCTGATCTTCTCAACTATCTAAGAAGTAAAGAGAAAATTTTACA 2191
 2108 Db AATATTGTTGCTATGCTGATCTCTCACTACCTAAGAGTAAAGAGAGTAAAGAGAGTTCACA 2167
 2192 QY GGAATTTGGAAGAGATTTTCAAGGAACACAATTTCAATTTGATTTTACCCCACTTCCATC 2251
 2168 Db GGAATTTGGAAGAGATTTTCAAGGAACACAATTTCAATTTGATTTTACCCCACTTCCAGGC 2227
 2252 QY ATCCAAATTTCCAGATGCTGTTCAAGAGAGTTCAGATACACCCGCTCCGATCAMA 2311
 2228 Db ATCCAAATTTCCAGATGCTGTTCAAGAGAGTTCAGATACACCCGCTCCGATCAGC 2287
 2312 QY TCTCAGGCTTCAATGGGAATTTCAATTTCACTCTCAAGATGAATTTGAATATGAATAACCAA 2371
 2288 Db TCTCAGGCTTCAATGGGAATTTCAATTTCACTCTCAAGATGAATTTGAATATGAATAACCAA 2347
 2372 QY AAAGGCTG-----GAAGAGAGGAGGACTTTGAATGTGCTTACATTTGAAGATCTTCTTT 2425
 2348 Db AGAGGCTGCGCAGAGAAGAGGAGGAGGATTTGAACGTGCTGACGTTTGAAGAGCTCTCTTT 2407
 2426 QY GCTTTGATATCAAGTTGCGCAAGGAATGAATTTCTGGAATTTAGTCTGTGTTTACA 2485
 2408 Db GCTTTGATATCAAGTTGCGCAAGGAATGAATTTCTGGAATTTAGTCTGTGTTTACA 2467
 2486 QY GAGACCTTGGCCGAGGAGAGTCTCTCACCACCGGAAAGTGGTGAAGATGTGACT 2545
 2468 Db GAGACCTTGGCCGAGGAGAGTCTCTCACCACCGGAAAGTGGTGAAGATGTGACT 2527
 2546 QY TTGGATTTGGCTCGAGATATCATGATGATTTCCAACTATGTTGTGAGGGGCAATGCGGCTC 2605
 2528 Db TTGGATTTGGCTCGAGATATCATGATGATTTCCAACTATGTTGTGAGGGGCAATGCGGCTC 2587
 2606 QY TGCTGTAAATGATGAGTGGCCCGGAAAGCTGTTTGAAGGATCTACACCATTTAAGAGTG 2665
 2588 Db TGCTGTAAATGATGAGTGGCCCGGAAAGCTGTTTGAAGGATCTACACCATTTAAGAGTG 2647
 2666 QY ATGCTGGTCAATGGAATATTAATGTTGGGAATCTTCTCACTTGGTGTGAATCTTTACC 2725
 2648 Db ATGCTGGTCAATGGAATATTAATGTTGGGAATCTTCTCACTTGGTGTGAATCTTTACC 2707
 2725 QY CTGGCATTTCCGGTTGATGCTAACTTCTACAACTGATTTCAAAATGATTTTAAATGATC 2785

2708 Db CTGGCATTTCCGTGCGACGCTAACTTCTATAAACTGATTCAGAGTGGATTTAAATGGAGC 2767
 2786 QY AGCCCAATTTATGCTACAGAGAAATATACATTTAATGCAATCCCTGCTGGCTTTTGACT 2845
 2768 Db AGCCCAATTTATGCTACAGAGAAATATACATTTAATGCAATCCCTGCTGGCTTTTGACT 2827
 2846 QY CAAAGAAAGCGGCATCTTCCCTAAATTTGACTTCTGTTTATAGATGTGAGTGGCAGATG 2905
 2828 Db CAAAGAAAGCGGCATCTTCCCTAAACCTGACTTCAATTTTATAGATGTGAGTGGCAGAGG 2887
 2906 QY CAAAGAAAGCGATGTATCAGATGTGGATGGCGCTGTTTCGGAAATGCTCTCACAACCTACC 2965
 2888 Db CAAAGAAAG-----CATGTATCAGAAATCATTCATCTCATCTACC 2924
 2966 QY AAAACAGGCGCACTTTTTCAGCAGAGAGATGGATTTGGGGCTACTCTCTCCGCGGCTCAGG 3025
 2925 Db AAAACAGGCGGCTTTCAGCAGAGAGGCGGCTCAGAGC-CCAGTCCGACACGCGCAGG 2983
 3026 QY TCGAAGATTTCTGAGAGGAAACAATTTAGTTTAAAGACTTCACTCCCTCCACTATCCCTAA 3085
 2984 Db T-GAAGATTTACAGAGAAAGAGTTAGCGAGGAGGCTTTGGACCCGCG-----CACCTAG 3038
 3086 QY CAGGCTGTAGATTACCAAAACAAGATTTATTTTATCATCTCAATAAGAAATCTATTATCAAC 3145
 3039 Db CAGGCTGTAGATCCGAGAGCCCAAGATTAGCTCGCCTCT--GAGGAAGCGCCCTACAGCG 3096
 3146 QY TGTCTGCTTCCACAGACTTTTCTCTAGAAGCGCTCTGCGTTTACTCTTGTTCCTTCAAGGGA 3205
 3097 Db TGTCTGCTTCCGCTGAGCTTTTCTCTAGATGCTGCTGCTGCTTCTC-----CAAAGTGA 3149
 3206 QY CTTTGTGTAATAATCAATCTCTGTCACAGGAGGAGGAGCTGATTAATGAATCTTTTATG 3265
 3150 Db CTTTGTGTAATAATCAATCTCTCTGTCACAGGCGGAGAGCCCAATTAATGAAGACTTTGTTG 3209
 3266 QY GAGCATTTGATCTGCATCCCAAGGCTTCTCAGGCGGCTTGTGATGATTTGTGTACCTGAAG 3325
 3210 Db TGAGCGCGCTTACCTTGGGGGCTTTTCCAGAGCTTGAAGGAAAGCCATGTATCTGAAA 3269
 3326 QY TACAGTATATTTCTGTAAATATACATAAAACAAA-----AGCATTTTGTCTAAGGAGAAC 3378
 3270 Db TATAGTATATTTCTGTAAATATACGTGAACCAACCAACCCGTTTGTCTAAGGAGAAC 3329
 3379 QY TAATATGATTTTCT--AAGTCTATGTTTAAATAATATGTAATTTTTCAGCTATTTAG 3436
 3330 Db TAATATGATTTTCTAATAAATCTATGTTTAAATACTATGTAATTTTTCATCTATTTAG 3389
 3437 QY TGAATATTTTATGCTGGGATTAATAATTTTCTACTACAGAAAAAATAAAAAAATAAAAA 3496
 3390 Db TGAATATTTTATGCTGGGATTAATAAATTTTCTACTGTAATAAATAAAAAAATAAAAAA 3449
 3497 QY AAAAA 3500
 3450 Db AAAAA 3453

Search completed: August 28, 2003, 01:02:16
 Job time : 16331.9 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 27, 2003, 15:17:42 ; Search time 8449.93 seconds
(without alignments)
10069.907 Million cell updates/sec

Title: US-09-919-408a-3
Perfect score: 3501
Sequence: 1 CGAGGCGCATCCGAGGCT.....AAAAAAAAAAAAAAAAAAAA 3501

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

```
EST.*
1: em_estba.*
2: em_esthum.*
3: em_estinu.*
4: em_estinu.*
5: em_estov.*
6: em_estpl.*
7: em_estro.*
8: em_hic.*
9: gb_estl.*
10: gb_est2.*
11: gb_hic.*
12: gb_est3.*
13: gb_est4.*
14: gb_est5.*
15: em_estfun.*
16: em_estom.*
17: em_gss_hum.*
18: em_gss_inv.*
19: em_gss_pln.*
20: em_gss_vrt.*
21: em_gss_fun.*
22: em_gss_man.*
23: em_gss_mus.*
24: em_gss_pro.*
25: em_gss_rod.*
26: em_gss_phg.*
27: em_gss_vrl.*
28: gb_gss1.*
29: gb_gss2.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3299.4	94.2	3647	11	BC036028
2	846.4	24.2	3761	11	AK045865
3	650.6	18.6	714	12	BI461248
4	644.8	18.4	648	13	BU681356

5	509.6	14.6	621	12	BI360262
6	486.2	13.9	765	9	AI323253
7	440	12.6	604	10	BF523018
8	437.8	12.5	460	9	AV713950
9	402.6	11.5	618	9	AA120050
10	400.4	11.4	455	12	BM147250
11	389.6	11.1	507	12	BM484050
12	382	10.9	510	14	CB220941
13	367.4	10.5	492	13	BK279530
14	358.4	10.2	500	13	BU696796
15	341.2	9.7	3616	11	AK080326
16	340.4	9.7	3312	11	AK004947
17	340.4	9.7	3687	11	AK079247
18	337.2	9.6	3822	11	BC050024
19	334.2	9.5	3687	11	AK076215
20	325.6	9.3	445	10	BE651447
21	318.6	9.1	3883	11	AK046795
22	318.6	9.1	5174	11	BC026713
23	314.4	9.0	522	10	BB644407
24	307.6	8.8	3887	11	AK047010
25	306.8	8.8	597	13	BU695104
26	299.6	8.6	646	9	AW772610
27	291.6	8.3	658	14	BY727686
28	283.4	8.1	659	9	AW057705
29	274.8	7.8	568	10	BE679272
30	249.2	7.1	515	13	BQ556177
31	242.6	6.9	451	14	CB787035
32	241	6.9	332	14	CB605696
33	240.4	6.9	394	10	BG145808
34	238.2	6.8	724	14	BY764070
35	237.2	6.8	640	14	BY749257
36	236.8	6.8	449	13	BY545844
37	236.6	6.8	660	10	BE047675
38	235.6	6.7	939	13	BX371699
39	234.8	6.7	964	14	CD301880
40	230.8	6.6	779	14	CA511669
41	230.4	6.6	276	10	BF176247
42	228.4	6.5	739	10	BF159205
43	228	6.5	480	9	AI150354
44	227.2	6.5	651	14	CA334084

ALIGNMENTS

RESULT 1	BC036028	BC036028	3647 bp	linear	HTC 23-SEP-2002
LOCUS	Homo sapiens	clone IMAGE:5272266	mrna		
DEFINITION	BC036028				
ACCESSION	BC036028				
VERSION	BC036028.1	GI:23273564			
KEYWORDS	HTC				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	1	(bases 1 to 3647)			
AUTHORS	Strausberg,R.				
TITLE	Direct Submission				
JOURNAL	Submitted (31-JUL-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA				
REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov				
COMMENT	Contact: MGC help desk Email: cgabs-r@mail.nih.gov Tissue Procurement: Miklos Palkovits, M.D., Ph.D. CDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki Toshiyuki and Piero Carninci (RIKEN) CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL) DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305				

Db 1765 AATTGTGTCACAAAGTACAAAAGCAATTTAGGTATGAAAGCCAGCTACAGATGGTACAGGT 1824
Qy 1800 GACCGGCTCTCAGATAATGAGTACTTCTAGTGTGATTTTCAGAGAAAT-ATGAAATATGATC 1858
Db 1825 GACCGGCTCTCAGATAATGAGTACTTCTAGTGTGATTTTCAGAGAAATGAAATATGATC 1884
Qy 1859 TCAATGGGAGTTTCCAAAGAGAAAATTTAGAGTTTGGGAAGGTACTAGGATCAGGTGCTT 1918
Db 1885 TCAATGGGAGTTTCCAAAGAGAAAATTTAGAGTTTGGGAAGGTACTAGGATCAGGTGCTT 1944
Qy 1919 TTGAAAGTGAATGAACGCAACAGCTTATGGAATTTAGCAAAACAGAGAGTCTCAATCCAGG 1978
Db 1945 TTGAAAGTGAATGAACGCAACAGCTTATGGAATTTAGCAAAACAGAGAGTCTCAATCCAGG 2004
Qy 1979 TTGCGGTCAAAATGCTCAAGAGAAAAGCAGACAGCTCTGAAAGAGAGGCACTCATGTCAG 2038
Db 2005 TTGCGGTCAAAATGCTCAAGAGAAAAGCAGACAGCTCTGAAAGAGAGGCACTCATGTCAG 2064
Qy 2039 AACTCAAGATGATGACCCAGCTGGGAAGCCACAGAGATATTTGCAACCTGCTGGGGCGT 2098
Db 2065 AACTCAAGATGATGACCCAGCTGGGAAGCCACAGAGATATTTGCAACCTGCTGGGGCGT 2124
Qy 2099 GCACACTGTGAGGACCAATTTACTTTGATTTTGAATCTGTTGCTATGTTGATCTTCTCA 2158
Db 2125 GCACACTGTGAGGACCAATTTACTTTGATTTTGAATCTGTTGCTATGTTGATCTTCTCA 2184
Qy 2159 ACTATCTAAGAGTAAAGAGAAAATTTACAGGACTTGGACAGAGATTTTCAAGGAAC 2218
Db 2185 ACTATCTAAGAGTAAAGAGAAAATTTACAGGACTTGGACAGAGATTTTCAAGGAAC 2244
Qy 2219 ACAATTTCACTTTTACCCCACTTTTCAATCACATCCAAATTCAG- 2264
Db 2245 ACAATTTCACTTTTACCCCACTTTTCAATCACATCCAAATTTCCAGTAAAGAAATGAG 2304
Qy 2265 - 2264
Db 2305 CTTTACAAAGCAACTGGAAGAAAAGAGATGTTGTAACGCTTACGGGACTCTCGGAA 2364
Qy 2265 - 2264
Db 2365 GATCTGTATTTATGTGAGGAAAAGTGGGCTGAGCTCAGAAACCAAGAAATGAGATCGATCA 2424
Qy 2267 TGCGTGGTTCAAGAGAGTTTCAGATACACCGGACTCGGATCAAAATCTCAGGCTTCATG 2326
Db 2425 TGCGTGGTTCAAGAGAGTTTCAGATACACCGGACTCGGATCAAAATCTCAGGCTTCATG 2484
Qy 2327 GGAATTCATTTCACTCTGAAGATGAAATTTGAATATGAAACCAAAAGGCTGGAAGAAG 2386
Db 2485 GGAATTCATTTCACTCTGAAGATGAAATTTGAATATGAAACCAAAAGGCTGGAAGAAG 2544
Qy 2387 AGGAGACTTGAATGTGCTTACATTTCAAGATCTTCTTTGCTTTGCAATCAAGTTGCCA 2446
Db 2545 AGGAGACTTGAATGTGCTTACATTTCAAGATCTTCTTTGCTTTGCAATCAAGTTGCCA 2604
Qy 2447 AAGGAATGGAATTTCTGGAATTTTAAAGTCTGTGTTCTACAGAGACTCGGCCAGGAACG 2506
Db 2605 AAGGAATGGAATTTCTGGAATTTTAAAGTCTGTGTTCTACAGAGACTCGGCCAGGAACG 2664
Qy 2507 TGCTTTGACCCACCGGAAAAGTGGTGAAGATATGTGACTTTGGAATTTGGCTCGAGATATCA 2566
Db 2665 TGCTTTGACCCACCGGAAAAGTGGTGAAGATATGTGACTTTGGAATTTGGCTCGAGATATCA 2724
Qy 2567 TGAGTGAATCCAACTATGTTGTGACGGGCAATCCCGTCTGCTGTAATAATGGATGCC 2626
Db 2725 TGAGTGAATCCAACTATGTTGTGACGGGCAATCCCGTCTGCTGTAATAATGGATGCC 2784
Qy 2627 CCGAAGCCTGTTTGAAGGATCTACACCAATTAAGATGATGTTGTCATATGGAATAT 2686
Db 2785 CCGAAGCCTGTTTGAAGGATCTACACCAATTAAGATGATGTTGTCATATGGAATAT 2844
Qy 2687 TACTGTGGGAAATCTTCTCACTTGGTGTGAATCCTTACCCTGGCATTCGGTGTGATGCTA 2746
Db 2845 TACTGTGGGAAATCTTCTCACTTGGTGTGAATCCTTACCCTGGCATTCGGTGTGATGCTA 2904

Qy 2747 ACTTCTACAAACTGATTCAAAATGGAATTTAAATGGATCAGCCATTTTATGCTACAGAAG 2806
Db 2805 ACTTCTACAAACTGATTCAAAATGGAATTTAAATGGATCAGCCATTTTATGCTACAGAAG 2964
Qy 2807 AAATATATCATTATAATCAATCTCTGCTGGGCTTTTGAATCAAGGAAACGGCCATCTTCC 2866
Db 2965 AAATATATCATTATAATCAATCTCTGCTGGGCTTTTGAATCAAGGAAACGGCCATCTTCC 3024
Qy 2867 CTAAATTTGACTTCTGTTTTAGGATGTGAGTGGCAGATGCAAGAAAGCGATGTATCAGA 2926
Db 3025 CTAAATTTGACTTCTGTTTTAGGATGTGAGTGGCAGATGCAAGAAAGCGATGTATCAGA 3084
Qy 2927 ATGTGAGTGGCGGTTTTCGGAATGTCTCACACTACCAAAACAGGCGACTTTTCAGCA 2986
Db 3085 ATGTGAGTGGCGGTTTTCGGAATGTCTCACACTACCAAAACAGGCGACTTTTCAGCA 3144
Qy 2987 GAGAGATGGAATTTGGGGCTACTCTCTCCGAGGCTCAGGTGCAAGATTCGTAGAGGAACA 3046
Db 3145 GAGAGATGGAATTTGGGGCTACTCTCTCCGAGGCTCAGGTGCAAGATTCGTAGAGGAACA 3204
Qy 3047 ATTTAGTTTTAAGACTTTCATCTCCCTCACTATCCCTAAACAGGCTGTAGATTAACAAAC 3106
Db 3205 ATTTAGTTTTAAGACTTTCATCTCCCTCACTATCCCTAAACAGGCTGTAGATTAACAAAC 3264
Qy 3107 AAGATTAATTTTCATCATACTAAAGAAATCTATTAATCAACTGCTGCTTCCAGGACTTTTC 3166
Db 3265 AAGATTAATTTTCATCATACTAAAGAAATCTATTAATCAACTGCTGCTTCCAGGACTTTTC 3324
Qy 3167 TCTAGAAAGCGCTGCTGCTTTTACTCTCTTTTCAAGGAGCTTTTGTAAATCAAAATCATC 3226
Db 3325 TCTAGAAAGCGCTGCTGCTTTTACTCTCTTTTCAAGGAGCTTTTGTAAATCAAAATCATC 3384
Qy 3227 CTGTCAAGGCGAGGAGCTGATAATGAATTTTATTGGAGCAATTCATCTGCAATCCAAAG 3286
Db 3385 CTGTCAAGGCGAGGAGCTGATAATGAATTTTATTGGAGCAATTCATCTGCAATCCAAAG 3444
Qy 3287 GCCTTCTCAGGCGGCTTGTGATGAAATTTGTGCTGAAAGTACAGTATATTTCTGTAAATA 3346
Db 3445 GCCTTCTCAGGCGGCTTGTGATGAAATTTGTGCTGAAAGTACAGTATATTTCTGTAAATA 3504
Qy 3347 CATATAACAAAGCAATTTTCTAAGGAGCACTTAATATGATTTTAAAGTCTATGTTTTA 3406
Db 3505 CATATAACAAAGCAATTTTCTAAGGAGCACTTAATATGATTTTAAAGTCTATGTTTTA 3564
Qy 3407 AAATAATATCTAAATTTTTCAGCTTATTAGTATATATTTTATGGGTGGGAATAAAATTT 3466
Db 3565 AAATAATATCTAAATTTTTCAGCTTATTAGTATATATTTTATGGGTGGGAATAAAATTT 3624
Qy 3467 CTACTACAGAAAAA 3489
Db 3625 CTACTACAGAAAAA 3647

RESULT 2
AK045865
LOCUS

DEFINITION
AK045865 3761 bp mRNA linear HTC 05-DEC-2002
Mus musculus adult male corpora quadrigemina cDNA, RIKEN
full-length enriched library, clone:B20315G04 product:FMS-like
tyrosine kinase 3, full insert sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AK045865
AK045865.1 GI:26337656
HTC; CAP trapper.
Mus musculus (house mouse)

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED

1
Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
10349636
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

QY	494	AGAGTGAAGCTACCAATTAACAATAATCTTTACAGTGTAGTATAGAAATACCTGCTTT	553
Db	487	AGAGCGAAGCCCAACTACAGTACTGTTACAGTGAATGAAGAGATACACAGCTGT	546
QY	554	ACACATTAAGAGAGACTTACTTTAGAAAAATGMAAACCCAGGACGCCCTGGTCTGCATAT	613
Db	547	ACGTGCTAAGAGAGACTTACTTTAGAGAGATGMAAACCAGGACGACTGCTCTGCATCT	606
QY	614	CTGAGAGCGTTCCAGAGCGGATCGTGGAAATGGGTCTTTGCAATTCACAGGGGGAAGCT	673
Db	607	CCGAGGCTGTTCCAGAGCCCACTGTGGAGTGGTCTCTGACGCTCCCAACAGGAAAGCT	666
QY	674	GTAAGAAGAAAGTCCAGCTGTTGTTAAAGAGGAAAGAGTCTTCAATGATTTATTTG	733
Db	667	GTAAGAAGAAAGCGCCCTGCTGTTGTTCAGAAAGGAGAAAGTACTTCAATGATTTGTCG	726
QY	734	GGACGGACATTAAGTGTCTGTGCCAGAAATGAACTGGCGAGGGAATGACACAGCGTGTCA	793
Db	727	GACAGACATCAGATGCTGTGTAGAAATGCACTGGCGCGGAATGACCAAGCTGTCA	786
QY	794	CAATAGATCTAAATCAAACTCTCAGACACATGCCCAATATTTCTTAAAGTAGGGG	853
Db	787	CCATAGATCTAAACCAAGCTCTCAGAGCACACTGCCAGTATTTCTGAAAGTGGGG	846
QY	854	AACCTTATGATAGGTCGAAGCTGTTTCATGTGAACATGGATTCGGGCTCAGCTGGG	913
Db	847	AACCTTGTGGATCAGGTGTAAAGGCCATCCATGTGAACCATGGATTCGGGCTCAGCTGGG	906
QY	914	AATTAGAAAAACAAGCACTCGAGGAGGCAACTCTTTGAGATGATGATACCTATTCAACA	973
Db	907	AGCTGGAAGACAAGCCCTGGAGGAGGACGACTCTTTGAGATGATGATACCTATTCAACA	966
QY	974	ACAGAACTATGATACGGATTCCTTTGCTTTGTTATCATCATGTGCAAGAAACGACACCG	1033
Db	967	ACAGAACCATGATTCGGATTCCTTTGCTTTGTTATCATCATGTGCAAGAAACGACACCG	1026
QY	1034	GATACACTACATGTTCTCTTCAAGCAATCCAGTCAATCAAGTGTGTTGTTACCAATCGTAG	1093
Db	1027	GATATTACACTGCTCTTCTCAAGCAATCCAGTCAATCAAGTGTGTTGTTACCAATCGTAG	1086
QY	1094	GAAAGGATTTATAAAGCTACCAATTCAGTGAAGATTTGAAATTCACCAATATGAAG	1153
Db	1087	AAAAGGTTTATAAAGCTACCAATTCAGTGAAGATTTGAAATTCACCAATATGAAG	1146
QY	1154	AGTTTGTGTTTCTCTCAGGTTTAAAGCTTACCCACAAATCAGATGTACGTGGACCTTCT	1213
Db	1147	AGTTCTGTTCTCAGTCAAGTTTAAAGGTTACCCACAAATCCGATCCGATGCGATCTTCT	1206
QY	1214	CTGAAATCATTTCTTGTGAGCAAAAGGCTTTGTAACGGATACAGCATATCCAAAGT	1273
Db	1207	CTCAAGCTCATTTCTTGTGAACAGAGAGCGCTGGAGGATGGGTACAGACTCTCTCATT	1266

RESULT 3
BI461248
LOCUS 603206574F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:5272266 5',
DEFINITION mRNA sequence.
ACCESSION BI461248
VERSION BI461248.1 GI:15251904
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 714)
NIH-MGC http://mhc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
AUTHORS Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgabos-@email.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.

CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki Toshiyuki and Piero Carninci (RIKEN)
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLML1687 row: c column: 19
High quality sequence stop: 709.
Location/Qualifiers
1. 714
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5272266"
/lab_host="DH10B"
/clone_lib="NIH_MGC_97"
/note="Organ: testis; Vector: pBluescriptR (modified pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI (gtcgag); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVN-3', size-selected for average insert size 2.2 kb and normalized to ROT 5. This is a primary library enriched for full-length clones and constructed using the Cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NIH/NHGRI, National Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT 182 a 171 c 191 g 170 t
ORIGIN

Query Match	18.6%	Score 650.6;	DB 12;	Length 714;
Best Local Similarity	97.8%	Pred. No. 1.7e-72;		
Matches 670;	Conservative 0;	Mismatches 14;	Indels 1;	Gaps 1;
QY	1	CGAGCGCATCCCGAGGGCTGGCGGGCGC-CCTGGGGACCCCGGCTCCGGAGGCCAT	59	
Db	30	CGAGCGCATCCCGAGGGCTGGCGGGCGC-CCTGGGGACCCCGGCTCCGGAGGCCAT	89	
QY	60	CGCGCGTGTGGCGGGCGACCGCGGCGCCTGCGCGTCTGCTGTTGTTTCTGCAATGAT	119	
Db	90	CGCGCGTGTGGCGGGCGACCGCGGCGCCTGCGCGTCTGCTGTTGTTTCTGCAATGAT	149	
QY	120	ATTTGGGACTATTACAAATCAAGATCTCCCTGTGATCAAGTGTGTTTAAATCAATCAATA	179	
Db	150	ATTTGGGACTATTACAAATCAAGATCTCCCTGTGATCAAGTGTGTTTAAATCAATCAATA	209	
QY	180	GAACAATGATTCATCAGTGGGAGTCAATCATATCCCATGGTATCAGAAATCCCGGA	239	
Db	210	GAACAATGATTCATCAGTGGGAGTCAATCATATCCCATGGTATCAGAAATCCCGGA	269	
QY	240	AGACTCGGGTGTGCTTGGAGCCCGAGCTCAGGGACAGTGTACGAAGCTGCCGCTGT	299	
Db	270	AGACTCGGGTGTGCTTGGAGCCCGAGCTCAGGGACAGTGTACGAAGCTGCCGCTGT	329	
QY	300	GGAAAGTGATGTATCTGCTTCCATCACACTGCAAGTGTGTCGATGCCCGGGAACAT	359	
Db	330	GGAAAGTGATGTATCTGCTTCCATCACACTGCAAGTGTGTCGATGCCCGGGAACAT	389	
QY	360	TTCCCTGCTCTCGGCTCTTTAAGACAGCTCCCTGAATGCCAGGCACATTTTGAATTACA	419	
Db	390	TTCCCTGCTCTCGGCTCTTTAAGACAGCTCCCTGAATGCCAGGCACATTTTGAATTACA	449	
QY	420	AAACAGAGGAGTGTGTTTCCATGGTCAATTTGAAATGACAGAAACCCCAAGCTGGAGAATA	479	
Db	450	AAACAGAGGAGTGTGTTTCCATGGTCAATTTGAAATGACAGAAACCCCAAGCTGGAGAATA	509	
QY	480	CCTACTTTTTTATTCAGAGTGAAGCTACCAATTAACAATATTTGTTTACAGTGAGTATAAG	539	
Db	510	CCTACTTTTTTATTCAGAGTGAAGCTACCAATTAACAATATTTGTTTACAGTGAGTATAAG	569	
QY	540	AAATACCTGCTTTACATTAAGAACCTTACTTTAGAAAAATGAAAAACCCAGGACGC	599	
Db	570	AAATACCTGCTTTACATTAAGAACCTTACTTTAGAAAAATGAAAAACCCAGGACGC	629	

```

Qy 600 CTTGGTCTGCATATCTCAGACCGTTCCAGACCGATCGTGAATGGTGTCTTTCGGATTC 659
Db 630 CTTGGTCTGCATATCTCAGACCGTTCCAGACCGATCGTGAATGGTGTCTTTCGGATTC 689
Qy 660 ACAGGGGGAAGCTGTAAAGAGAA 684
Db 690 ACAGGGGGAAGCTGTAAAGAGCA 714

RESULT 4
BU681356/c
LOCUS BU681356 648 bp mRNA linear EST 07-OCT-2002
DEFINITION UI-CF-EC1-abv-d-14-0-UI-sl UI-CF-EC1 Homo sapiens cDNA clone
ACCESSION BU681356
VERSION UI-CF-EC1-abv-d-14-0-UI 3', mRNA sequence.
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 648)
Bonaldo, M.F., Lennon, G. and Soares, M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
97044477
8889548
Contact: McCray, PB
McCray Lab
University of Iowa
2024 University of Iowa Med Labs, Iowa City, IA 52242, USA
Tel: 319 356 4866
Fax: 319 356 7171
Email: paul-mccray@uiowa.edu
Tissue Procurement: Dr. M. J. Welsh, University of Iowa
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com) or from Open Biosystems
(www.openbiosystems.com).
Seq primer: M13 FORWARD
POLYA=Yes.

FEATURES
source
1..648
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-CF-EC1-abv-d-14-0-UI"
/tissue_type="Lung"
/dev_stage="adult and fetal"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-CF-EC1"
/notes="Organ: Lung; Vector: pT7T3-Pac (Pharmacia) with a
modified polylinker; Site 1: EcoR I; Site 2: Not I;
UI-CF-EC1 is a normalized cDNA library containing the
day 64, day 87, week 19 and week 42. The library was
constructed according to Bonaldo, Lennon and Soares,
Genome Research, 6:791-806, 1996. First strand cDNA
synthesis was primed with an oligo-dT primer containing a
Not I site. Double stranded cDNA was ligated to an EcoR I
adaptor, digested with Not I, and cloned directionally
into pT7T3-Pac vector. The oligonucleotide used to prime
the synthesis of first-strand cDNA contains a library tag
sequence that is located between the Not I site and the
(dT)18 tail. The sequence tag for this library is
AAGTGCCTTAC.
TAG_LIB=UI-CF-EC1
TAG_TISSUE=Normal Lung Epithelial Cells Tissue nos 369-371
and 380-383
TAG_SEQ=AAGTGCCTTAC"

```

```

BASE COUNT 196 a 126 c 120 g 206 t
ORIGIN
Query Match 18.4%; Score 644.8; DB 13; Length 648;
Best Local Similarity 99.7%; Pred. No. 9.6e-72;
Matches 646; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 2844 CTCAGGAAACGGCCATCTTCCCTAATTTGACTTCGTTTTAGGATGTACAGTGGCAGA 2903
Db 648 CTCAGGAAACGGCCATCTTCCCTAATTTGACTTCGTTTTAGGATGTACAGTGGCAGA 589
Qy 2904 TGCAGAAAGACGATGTATCAGATGTGGATGGCCGTGTTTCGGAATGCTCTCACACCTA 2963
Db 588 TGCAGAAAGACGATGTATCAGATGTGGATGGCCGTGTTTCGGAATGCTCTCACACCTA 529
Qy 2964 CCAAACAGCGACCTTTTCAGCAGAGAGATGGATTTGGGGCTACTCTCTCCGCGAGGCTCA 3023
Db 528 CCAAACAGCGACCTTTTCAGCAGAGAGATGGATTTGGGGCTACTCTCTCCGCGAGGCTCA 469
Qy 3024 GGTGGAAGATTCGTAGAGGAACTAATTTAGTTTAAAGACTTTCATCCCTCCACCTATCCCT 3083
Db 468 GGTGGAAGATTCGTAGAGGAACTAATTTAGTTTAAAGACTTTCATCCCTCCACCTATCCCT 409
Qy 3084 AACGGCTGTAGATTACCAAAACAGATTAATTTTCATCATACTAAAGAAAATCTATTATCA 3143
Db 408 AACGGCTGTAGATTACCAAAACAGATTAATTTTCATCATACTAAAGAAAATCTATTATCA 349
Qy 3144 ACTGCTCTTCCACGAGACTTTTCTCTAGAAGCCGTCTGCGTTTACTCTGTTTTCAAAAGG 3203
Db 348 ACTGCTCTTCCACGAGACTTTTCTCTAGAAGCCGTCTGCGTTTACTCTGTTTTCAAAAGG 289
Qy 3204 GACTTTTGTAAATCAAAATCATCTCTGTCACAAGCAGGAGGAGCTGATAATGAACCTTAT 3263
Db 288 GACTTTTGTAAATCAAAATCATCTCTGTCACAAGCAGGAGGAGCTGATAATGAACCTTAT 229
Qy 3264 TGGAGCATTTGATCTGCATCCAAAGCCCTTCTCAGCGCCGCTTGTAGTGAATTTGTACCTGA 3323
Db 228 TGGAGCATTTGATCTGCATCCAAAGCCCTTCTCAGCGCCGCTTGTAGTGAATTTGTACCTGA 169
Qy 3324 AGTACAGTATATCTTCTGTAATACATAAAACAAAGCATTTTGTCTAGGGAAGCTAATA 3383
Db 168 AGTACAGTATATCTTCTGTAATACATAAAACAAAGCATTTTGTCTAGGGAAGCTAATA 109
Qy 3384 TGATTTTAAAGTCTATGTTTAAATAATATGTAATTTTTCAGCTATTTAGTGATATA 3443
Db 108 TGATTTTAAAGTCTATGTTTAAATAATATGTAATTTTTCAGCTATTTAGTGATATA 49
Qy 3444 TTTATCGGTGGGAATATAATTTCTACTACAGAAAAAATAAAAAA 3491
Db 48 TTTATCGGTGGGAATATAATTTCTACTACAGAAAAAATAAAAAA 1

RESULT 5
BI360262 621 bp mRNA linear EST 01-AUG-2001
LOCUS BI360262
DEFINITION 387099 MARC 2P1G Sus scrofa cDNA 5', mRNA sequence.
ACCESSION BI360262
VERSION BI360262.1 GI:15056290
KEYWORDS EST.
SOURCE Sus scrofa (pig)
ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
1 (bases 1 to 621)
Fahrenkrug, S.C., Smith, T.P.L.; Freking, B.A., Cho, J., White, J.,
Vallet, J., Wise, T., Rohrer, G.A., Pertea, G., Sultana, R., Quackenbush,
J. and Keele, J.W.
Porcine gene discovery by normalized cDNA-library sequencing and
EST cluster assembly
Mamm. Genome 13 (8), 475-478 (2002)
22213789
PUBMED 12226715
Contact: Smith TPL

```



```

Qy 1642 TTAACCTCTCAGGCGCCCTTCCCTTTTCATCAAGACAAACATCTCATCTATGCAACAAATT 1701
Db 61 TTAACCTCACCAGGCGCCCTTCCCTTTTCATCAAGACAAACATCTCATCTATGCGACCAATT 120
Qy 1702 GGTGTTTCTCCTCTTCATCTATGCTGTTTAAACCTGCTTAATTTCTCAACAGTACAAAAG 1761
Db 121 GGGCTCTGTCTCCCTCTCATCTATGCTGTTTAAACCTGCTTAATTTCTCAACAGTACAAAAG 180
Qy 1762 CAATTAGGTATGAAGGACGCTACAGATGTCAGAGTGCACGGCTCTCCAGATAATGAG 1821
Db 181 CAATTAGGTACGAGAGTCACTGACAGATGTCAGAGTGCACGGCTCTCCAGATAATGAG 240
Qy 1822 TACTTCTACGTTGATTTTCAGAGAATATGAATATGATCTCAATGGAGTTCCTCAAGAGAA 1881
Db 241 TACTTCTACGTTGATTTTCAGAGAATATGAATATGATCTCAATGGAGTTCCTCAAGAGAA 300
Qy 1882 AATTAGGTTTGGAGGTTACTAGGATCAGGTCGTTTGGAGAGTATGATGAACGCCACG 1941
Db 301 AACTTAGGTTTGGAGGTTACTAGGATCAGGTCGTTTGGAGAGTATGATGAACGCCACG 360
Qy 1942 GCTTATGGAATTAGCAAAACAGGAGTCTCAATCCAGGTTGCCGTCAAAATGCTGAAAGAA 2001
Db 361 GCCTATGGCAATTAGTAAACCGGAGTCTCAATTCAGGTGGCGTCAAGATCTAAAGAG 420
Qy 2002 AAGCAGACAGCTCTGAAAGAGAGGCACTATGTCAGAACTCAAGATGATGACCCAGCTG 2061
Db 421 AAGCTGACAGCTCTGAAAGAGAGGCACTATGTCAGGAGCTCAAAATGATGACCCAGCTT 480
Qy 2062 GGAAGCCACAGATATGTAACCTGCTGGGGCGTGCACATGTCAGGACCAATTTAC 2121
Db 481 GGAACATGACCACTGCTGGAGTTGTTGGGGCTTCCACATGTCAGGCGCCAGTGCAC 540
Qy 2122 TTGATTTTGAATATGTTGCTATGCTGATCTTCAACTATCTAAGAAGTAAAGAGAA 2181
Db 541 TGGATTTTGAATATGTTGCTATGAGACCTCTCACTACTCAAGATGATGAAGAGAG 600
Qy 2182 AATTTCACAGGACTTGACAGAGATTTCAAGAAACAAATTTTCAGTTTTCACCCACT 2241
Db 601 AATGTTTCACAGGACTTGACAGAGATTTTAAAGACCAATAA-TTCAGTTTTCACCTTACT 659
Qy 2242 TTCAATCAGATCAAAATTCAGAGTCCCTGTTTCAGAGAGT 2285
Db 660 TTCCACAGATCAAAATTCAGAGTCCCTGTTTCAGAGAGT 703

```

```

RESULT 7
BF523018 604 bp mRNA linear EST 11-DEC-2000.
DEFINITION UI-R-C2p-rg-f-10-0-UI.r1 UI-R-C2p Rattus norvegicus cDNA clone
ACCESSION BF523018
VERSION BF523018.1 GI:11631033
KEYWORDS EST.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 604)
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene
discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
PUBMED 8889548
COMMENT Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@uiowa.edu
cDNA Library Preparation: M.B. Soares Lab Clone distribution:

```

clones will be available through Research Genetics (www.resgen.com)
 This clone is also available through the I.M.A.G.E. Consortium at
 LNL (info@image.llnl.gov). IMAGE ID= 1793119
 Seq primer: M13 Forward.

FEATURES

Location/Qualifiers
 1..604
 /organism="Rattus norvegicus"
 /mol_type="mRNA"
 /strain="Sprague-Dawley"
 /db_xref="taxon:10116"
 /clone="UI-R-C2p-rg-f-10-0-UI"
 /dev_stage="adult"
 /lab_host="DH10B (Life Technologies)"
 /clone_lib="UI-R-C2p"
 /note="vector: p7T3D-Pac (Pharmacia) with a modified
 polylinker; Site 1: Not I; Site 2: Eco RI; The UI-R-C2p
 library is a subtracted library derived from the UI-R-C1
 library, which is a subtracted library derived from the
 UI-R-C0 library. The UI-R-C0 library consisted of a
 mixture of individually tagged normalized libraries
 constructed from rat placenta, adult lung, brain, liver,
 kidney, heart, spleen, ovary, muscle, 8, 12 and 18-day
 embryo. The tag is a string of 3-5 nucleotides present
 between the Not I site and the oligo-dT track which allows
 identification of the library of origin of a clone within
 the mixture. The subtracted library (UI-R-C2p) was
 constructed as follows: PCR amplified cDNA inserts from
 UI-R-C1 clones from which 3' ESTs had been derived was
 used as a driver in a hybridization with the UI-R-C1
 library in the form of single-stranded circles. The
 remaining single-stranded circles (subtracted library) was
 purified by hydroxyapatite column chromatography,
 converted to double-stranded circles and electroporated
 into DH10B bacteria (Life Technologies) to generate the
 UI-R-C2p library. This procedure has been previously
 described (Bonaldo, Lennon and Soares, Genome Research 6:
 791-806, 1996)"

BASE COUNT 166 a 132 c 167 g 139 t
 ORIGIN

Query Match 12.6%; Score 440; DB 10; Length 604;
 Best Local Similarity 85.4%; Pred. No. 4.9e-46;
 Matches 504; Conservative 0; Mismatches 80; Indels 6; Gaps 1;

Qy 2016 TGAAGAGAGGCACTCATGTCAAGAACTGATGACCCAGCTGGAGCCACGAGAA 2075
 Db 9 TGAAGAGAGGCGCTCATGGCTGAGCTCAAAATGATGACCCAGCTGGGACCAATGACNA 68
 Qy 2076 TATTGTGAACCTGCTGGGGCGTGCACACTGTCAAGACCAATTTTACTTTTGAATA 2135
 Db 69 CATCGTGAACCTGCTGGGGCGATGCACACTGTCAAGGCCAGTGTACTTTTGAATA 128
 Qy 2136 CTGTTGCTATGGTATCTTCTCAACTATCTAAGAGTAAAGAGAAAAATTCACAGAC 2195
 Db 129 TTGTTGTCATGGTGACCTGCTCACTACCTAAGAGCAAAAGAGAAAAAGTTTCACAGAC 188
 Qy 2196 TTGACACAGATTTTCAAGGAACACAAATTTTACCTTTTACCCACTTTCCCAATCACATCC 2255
 Db 189 GTGACACAGATTTTAAAGGAACATAATTTTACGTTTTTACCCACTTTCCAGTCAATTC 248
 Qy 2256 AAATTCAGCATGCTGTTTCAAGAGAGTTCAGATACACCCGACCTCGGATCAAAATCTC 2315
 Db 249 AAATTCAGTATGCCGGTTTCAAGAGAGTTCAGATATACCCGCGCTGGATCAGGTCTC 308
 Qy 2316 AGGGCTTCATGGGAATTCATTTTCACTCTGAAGATGAAATGAAATGAAACCAAAAAAG 2375
 Db 309 AGGGTTCAATGGGAATTCATTTTCAATTTGAAAGATGAGATGAGTATGAAACCAAGAG 368
 Qy 2376 GCTGGAAGAGAGAGAG-----GACTTGAATGTCTTACATTTGAAGATCTTCTTTGCTT 2429
 Db 369 GCTGGAAGAGAGAGAGAGATTTTGAACGCTGACGTTTGAAGACCTCTCTTTGTTT 428
 Qy 2430 TGCATATCAAGTTGCCAAAGGAATGGAATTTCTGGAATTTAAGTCGTGTGTTCACAGAGA 2489

[illegible]

KEYWORDS EST.
SOURCE Sus scrofa (pig)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.

REFERENCE 1 (bases 1 to 507)
AUTHORS Fahrrenkrug S.C., Smith T.P.L., Freking B.A., Cho J., White J., Vallet J., Wise T., Rohrer G.A., Pertea G., Sultana R., Queckenbush J., and Keele J.W.

TITLE Porcine gene discovery by normalized cDNA-library sequencing and EST cluster assembly
EST cluster assembly
Mamm. Genome 13 (8), 475-478 (2002)

JOURNAL 22213789
MEDLINE 12226715
PUBMED
COMMENT Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and alt. trimmed with phred v0.980904.e. Vector identified by cross_match with the -minscore 18 and -minmatch 12 options.
PCR Primers
FORWARD: AGGAACAGCTATGACCAT
BACKWARD: GTTTCACAGTCACGACG
Plate: 2 row: D column: 23
Seq primer: ATTAGTGACACTATAG.

FEATURES
Location/Qualifiers
source 1..507
/organism="Sus scrofa"
/mol_type="mRNA"
/db_xref="taxon:9823"
/tissue_type="pooled"
/lab_host="DH108"
/clone_lib="MARC 2P1G"
/notes="Vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI; Library made from pooled tissue from testis, ovary, endometrium, hypothalamus, pituitary, and placenta."
BASE COUNT 144 a 132 c 118 g 112 t 1 others
ORIGIN
Query Match 11.1%; Score 389.6; DB 12; Length 507;
Best Local Similarity 88.0%; Pred. No. 1.1e-39;
Matches 447; Conservative 0; Mismatches 59; Indels 2; Gaps 2;

QY 1194 CAGATGCTGGACCTTCTCTCGAAATCATTTCTTGTCGACCAAGGCTTGTATAA 1253
DB 1 CAGGTGACGTGGACCTTCTCTCGAAA-CATTTCTTGTCGACCAAGGCTGACGCA 59

QY 1254 CGGATACAGCATATCCAAAGTTTTCGAATCATAGCACCAGGAGAGATATATATCCA 1313
DB 60 CGGGTACAGCATATCTAAGTTTTCGAACCATACACAGCAGGACATACATCTCCA 119

QY 1314 TGCA-GAAATGATGATGCCAATTTACCAAAATGTTACGCTGGAATATAAGAGGAAC 1372
DB 120 TGCAGAAAATGACGACGCGCCAGTTCACAAAGATGTTACGCTGGAATATAAGAGGAAC 179

QY 1373 CTCAGTGTCCGAGAGCATCGGAAGTCAGCGCTCTGTTCTCGGATGGATACCCAT 1432
DB 180 CGCAAGTGTGGCCGAGGATCGGCAAGTCAGCGCTCTGTTCTGATGGTTACCCGT 239

QY 1433 TACCATTCTGGACCTGGAAGAAGTGTTCAGACAAGTCTCCCACTGCACAGAAGAGATCA 1492
DB 240 TACCGTCTTGGACCTGGAAGAAGTGTTCAGACAAGTCTCCCACTGCACAGAAGAGATCA 299

QY 1493 CAGAAAGAGTCTGGAATAGAAAGGCTTAACAGAAAAGTGTTCGACAGTGGGTTCGAGCA 1552
DB 300 CGGAAGGATCTGGAATAAAAGGCCCAACAGGAAAGTATTTTGACAGTGGATTTCCAGCA 359

QY 1553 GTACTCTAAACATGAGTGAAGCCATAAAGGGTCTCTGGTCAAGTGTGTGCATACAAT 1612
DB 360 GCACTCTGAACATGACGAGGCTGTCAAAAGGGTCTCTGGTGAAGTGTGTGCTGCTACAAT 419

QY 1613 CCCTGGCACATCTTGAGACGATCCTTTTAAACTCTCGAGGCCCTTCCTCCCTTCATCC 1672
DB 420 CCTGGGCAGCTCTTGGAACGATCCTCTTAACTCACCAGGCCCTTCCTCCCTTCATCC 479

QY 1673 AAGACACATCTCTATCTATGCAACAAT 1700
DB 480 AGGACACATCTCTATCTATGCAACAAT 507

RESULT 12
CB220941
LOCUS
DEFINITION lAbo28C03 Bos taurus Abomasum #1 library Bos taurus cDNA similar to FLT3 receptor tyrosine kinase - GO terms: protein tyrosine kinase (0004713); receptor (0004872);, mRNA sequence.
ACCESSION CB220941
VERSION CB220941.1 GI:28291455
KEYWORDS EST.
SOURCE Bos taurus (cow)
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Bovinae; Bos.
REFERENCE 1 (bases 1 to 510)
AUTHORS Hansen, C., Fu, A., Meng, Y., Li, C., Okine, E., Sensen, C.W., Gordon, P.M.K. and Moore, S.S.
TITLE Gene Expression Profiling of the Bovine Gastrointestinal Tract
JOURNAL Unpublished
COMMENT Contact: Dr. Stephen Moore
Beef Genomics Laboratory
Dept of AFNS, University of Alberta
410 Agri/For, Dept of AFNS, U of A, Edmonton, AB, T6G 2P5, Canada
Tel: 780 492 0169
Fax: 780 492 4265
Email: stephen.moore@ualberta.ca
Insert Length: 510 Std Error: 0.00
POLYA=NO. Location/Qualifiers
source 1..510
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/tissue_type="Smooth muscle"
/cell_type="Epithelial"
/dev_stage="Young adult"
/lab_host="XLI-BlueWRF"-strain"
/clone_lib="Bos taurus Abomasum #1 library"
/notes="Organ: Abomasum; Vector: Uni-2ZAPXR; Site 1: EcoR I; Site 2: Xho I"
BASE COUNT 142 a 120 c 107 g 141 t
ORIGIN
Query Match 10.9%; Score 382; DB 14; Length 510;
Best Local Similarity 85.7%; Pred. No. 1e-38;
Matches 437; Conservative 0; Mismatches 70; Indels 3; Gaps 1;

QY 2644 GCATCTACACCAATTAAGTGTGATGCTGGTCATATGATATATCTGCGGAATCTTC 2703
DB 1 GCATCTACACCAATTAAGTGTGATGCTGGTCATATGATATATCTGCGGAATATTC 60

QY 2704 TCATTGGTGTGAATCCTTACCTCGGTTGATGCTAACTTCTCAAACTGATT 2763
DB 61 TGGTGTGTGTCATCTTACCTCGGATTCGCGTGTGCTAACTTCTCAAACTCATC 120

QY 2764 CAAATGGATTTAAATATGGATCAGCCATTTTATGCTACAGAGAAATATATATTAATG 2823
DB 121 CAGAGTGGGTTCAGATGACAGCCATTTTATGCCAGCAAGAAATATATATTAATG 180

QY 2824 CAATCTGCTGGCTTTTGACTCAAGGAAACGGCCATCTCTCCCTAAATTTGACTTCGTTT 2883
DB 181 CAATCTGCTGGCTTTTGACTCAAGGAAACGGCCATCTCTCCCTAACTGACTTCATTT 240

```

QY 2884 TTAGGATGTGAGTGGCAGATGCAGAAAGCGATGATCAGAAATGGATGGCGGTGT 2943
Db 241 TTAGGGGTGTGAGTGGCAGATGCAGAAAGCGATGATCAGAAATGGATGGCGGTGT 300
QY 2944 TCGGAATGTCTCTACACCTACCAAAACAGGCGACCTTTTCAGCAGAGAGATGGATTTGGGG 3003
Db 301 TCAGAGGCTCTCTGTTTACCAAAACAGGCGACCTTTTCAGCAGAGAGATGGATTTGGGG 360
QY 3004 CTACTCTCTCCGAGGCTCAGGTGCAAGATTCGATAGGAAACAATTTAGTTTAAAGACT 3063
Db 361 CTACCCCTCTCTCAGGTTACAGTGAAGATTCCTTAGAGAAACGATTTAGTCTTAAGGACC 420
QY 3064 TCATCCCTCCACCTATCCCTAACAGG---CTGTAGATTACCAAAACAGATTAATTTTCAT 3120
Db 421 TCATCCCTCCCTCCCTACCTTTAAACAGGCTGTGTAGATTACCAAAATAAGATTAATTTTCAT 480
QY 3121 CACTAAAAGAAATCTATTATCAACTGCTG 3150
Db 481 CACTAAGTGAATATATTCTCACTGCTG 510

RESULT 13
BX279530
LOCUS
DEFINITION BX279530 NCI CGAP Kid11 Homo sapiens cDNA clone IMAGE9998A067500 ;
IMAGE:3033869, mRNA sequence.
ACCESSION BX279530
VERSION BX279530.1 GI:28612110
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Ebert, L., Heil, O., Hennig, S., Neubert, P., Partsch, E., Peters, M.,
Radelof, U., Schneider, D. and Korn, B.
Human UnigeneSet - RZPD3
Unpublished
Contact: Ina Rolfs
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany
RZPD; IMAGE998A067500.
RZPDLIB; I.M.A.G.E. cDNA Clone Collection;
http://www.rzpd.de/CloneCards/cgi-
bin/showLib.pl.cgi?response?libNo=972 Contact: Ina Rolfs
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Heubnerweg 6, D-14059 Berlin, Germany
Tel: +49 30 32639 101
Fax: +49 30 32639 111
www.rzpd.de
This clone is available royalty-free from RZPD;
contact RZPD (clone@rzpd.de) for further information. Seq primer:
M13r, Primer sequence: TTTCAACAGAAACAGCTATGAC.
FEATURES
Location/Qualifiers
1..492
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE998A067500" ; IMAGE:3033869"
/lab_host="DH10B"
/clone_lib="NCI CGAP Kid11"
/note="Organ: kidney; Vector: pT73D-Pac (Pharmacia) with
a modified polylinker; Site 1: Not 1; Site 2: Eco RI;
Plasmid DNA from the normalized library NCI CGAP Kid3 was
prepared, and ss circles were made in vitro. Following HAP
purification, this DNA was used as tracer in a subtractive
hybridization reaction. The driver was PCR-amplified cDNAs
from a pool of 5,000 clones made from the same library
(cloneIDs 1322376-1323911, 1456007-1456775, and
1500552-1502855). Subtraction by Bento Soares and M.
Fatima Bonaldo. "
BASE COUNT 147 a 102 c 92 g 149 t 2 others

```

```

ORIGIN
Query Match 10.5%; Score 367.4; DB 13; Length 492;
Best Local Similarity 99.2%; Pred. No. 6.9e-37;
Matches 368; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2916 GATGATCAGAATGTGATGGCGGTGTTTCGGGAATGTCCTCACACCTACCAAAACAGCG 2975
Db 122 GATGATCAGAATGTGATGGCGGTGTTTCGGGAATGTCCTCACACCTACCAAAACAGCG 181
QY 2976 ACCTTTTCAGCAGAGAGATGATTTGGGGCTACTCTCTCCGAGGCTCAGGTTCGAAGATTC 3035
Db 182 ACCTTTTCAGCAGAGAGATGATTTGGGGCTACTCTCTCCGAGGCTCAGGTTCGAAGATTC 241
QY 3036 GTAGAGAAACAATTTAGTTTAAAGGACTTCATCCCTCCACCTATCCCTAACAGGCTGTAG 3095
Db 242 GTAGAGAAACAATTTAGTTTAAAGGACTTCATCCCTCCACCTATCCCTAACAGGCTGTAG 301
QY 3096 ATTACCAAAACAGATTAATTTTCATCACTAAAGAAATCTATTATCAACTGCTGCTTCA 3155
Db 302 ATTACCAAAACAGATTAATTTTCATCACTAAAGAAATCTATTATCAACTGCTGCTTCA 361
QY 3156 CCAGACTTTTCTAGAACGCGTCTGCGTTTACTCTTGTGTTTCAAAGGACTTTTGTAAA 3215
Db 362 CCAGACTTTTCTAGAACGCGTCTGCGTTTACTCTTGTGTTTCAAAGGACTTTTGTAAA 421
QY 3216 ATCAATATCATCTGTCTCAAGGAGGAGGAGCTGATTAATGAACCTTTTATGGAGCATTTGAT 3275
Db 422 ATCAATATCATCTGTCTCAAGGAGGAGGAGCTGATTAATGAACCTTTTATGGAGCATTTGAT 481
QY 3276 CTGCATCCAG 3286
Db 482 CTGCATCCNAG 492

RESULT 14
BX696796
LOCUS
DEFINITION BU696796 500 bp mRNA linear EST 09-OCT-2002
L2in11124T7 Hematopoietic Stem Cell Subtracted Library Mus
musculus cDNA 5' similar to human flk-2, mRNA sequence.
ACCESSION BU696796
VERSION BU696796.1 GI:23607570
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Phillips, R.L., Ernst, R.E., Brunk, B.P., Ivanova, N., Mahan, M.A.,
Deanehan, J.K., Moore, K.A., Overton, G.C. and Lemischka, I.R.
The genetic program of hematopoietic stem cells
Science 288 (5471), 1635-1640 (2000)
20295303
PUBMED 10834841
COMMENT
Contact: Lemischka, Ihor R.
Department of Molecular Biology
Princeton University
Lewis Thomas Laboratory, Washington Road, Princeton, NJ 08544, USA
Tel: 609 258 2838
Fax: 609 258 2759
Email: ilemischka@molbio.princeton.edu
These ESTs are derived from a subtracted cDNA library enriched for
gene products expressed in day 14-14.5 fetal liver hematopoietic
stem cells defined as Lineageneg/lo, AA4.lpos, ckltpos, Ly6A/E
(sca-1)pos
Seq primer: M13Reverse or T7.
Location/Qualifiers
1..500
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/tissue_type="Hematopoietic"
FEATURES
source

```


URL: <http://genome.gsc.riken.go.jp/>, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)

COMMENT
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Tissues were provided by Dr. John Todd (Dept. of Medical Genetics
Wellcome Trust Centre for Molecular Mechanisms in Disease Wellcome
Trust/MRC building Addenbrookes Hospital Cambridge) whose
assistance we gratefully acknowledge.
Please visit our web site for further details.
URL: <http://genome.gsc.riken.go.jp/>
URL: <http://fantom.gsc.riken.go.jp/>
Location/Qualifiers

FEATURES

source

1. 3616
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM DB:A630054K21"
/db_xref="taxon:10090"
/clone="A630054K21"
/tissue_type="thymus"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="3 days neonate"
76. 3014
/notes="colony stimulating factor 1 receptor
(MGDI|GI:1339758, GB|NM_007779, evidence: BLASTN, 98%,
match=3661)
putative"

BASE COUNT 854 a 998 c .972 g 792 t
ORIGIN

Query Match 9.7%; Score 341.2; DB 11; Length 3616;
Best Local Similarity 58.5%; Pred. No. 4.8e-34;
Matches 674; Conservative 0; Mismatches 463; Indels 15; Gaps 4;
QY 1715 TCTTCAATGTGCTTTAAACCTGCTAAATTTGTCAAGATACAAAAGCAATTTAGGTATG 1774
DB 1648 TCATGTTTCTGCTGGTGCTACTGCTGTGCTCTTGTACAGTACAAAGCAAGCCGA 1707
QY 1775 AAAGCCAGCTACAGATGTACAGGTGACCGGCTCTCAGATATAGTACTTCTACGTTG 1834
DB 1708 AGTACCAGGTGCGCTGGAAGATCATCGAGAGATACGAAGGCAATAGTACACCTTCATTG 1767
QY 1835 ATTTCAGAGATATGATATGATCTCAATGGAGTTTCCAGAGAAAATTTAGAGTTTG 1894
DB 1768 ACCCTACTCAGTTGCCCTACAATGAGAAGTGGAGTTCCCTCGGAACAACCTGCAGTTG 1827
QY 1895 GGAAGGTACTAGGATCAGGTGCTTTTGGAAAAGTGATGAACGCAACAGCTTATGGAATTA 1954
DB 1828 GTAAGACTTAGAGCGCGTGCTTTGGNAGGTGGTGGAGGCTACAGCCTTTGGTCTGG 1887
QY 1955 GCAAAAACAGAGTCTCAATCCAGGTTGCCGTCAAAATGTCTGAAGAAAAGACAGAGCT 2014
DB 1888 GCAAGAAGATGCAAGTGTGAAGTGGCTGTGAAGATGCTTAAAGTCCACGGCTCATGCTG 1947
QY 2015 CTGAAGAGAGGACATCATGTCAAGATCAAGATGATGACCCAGCTGGGAGCCAGAGA 2074
DB 1948 ATGAGAAGAGGCGCCCTGATGTCAAGTGAAGATCATGAGTCACTCGGAGCAGCAGAGA 2007
QY 2075 ATATTGTGAACCTGCTGGGGCGGTGCACACTGTGAGGACCAATTTACTTCAATTTTGAAT 2134
DB 2008 ATATAGTCAACCTTTGGAGCCCTGACTCAGGAGGACCTGCTCGGTGCTCACTGAAT 2067
QY 2135 ACTGTGCTATGGTGATCTTCTCAACTATCTAAAGAAAGTAAAGAGAAAATTTACAGGA 2194
DB 2068 ACTGCTGCTATGGAGACCTACTCAACTTCTCCGAAG-GAAGCCGAGGCTATGCTAGGA 2126
QY 2195 CTGGAACAGAGATTTCAAGGAACAATTTCAATTTTACCCCACTTTCCCAATCATC 2254
DB 2127 CCCAGCCTGAGTCTGCTGAGGACTCCGAGGGAGACTCCAGC-----TACAAGAACATC 2180

Search completed: August 28, 2003, 05:42:18
Job time : 8457.93 secs

QY 2255 CAAATTCACGATGCCCTGGTTCAAGAGAACTTCAGATACACCCGGACTCGGATCAAAATCT 2314
DB 2181 CACCTGGAGAGAAATATGTGCGCAGGACAGAGTGGCTTCTCCAGTCAAGGTGTAGACACC 2240
QY 2315 CAGGGCTTCATGGGAATTCATTTCACTCTCAAGATGAATTAATGATATGAACCAAAAA 2374
DB 2241 TAGCTGGAGATGAGGCTGTCTCGACT-----TCTTCAAGTGACTCTCTTTAAGCAAG 2295
QY 2375 GGCTGGAAGAAGAGGAGGACTTGAATGTGTATCATTTGAAGATCTTCTTTGCTTTGTCAT 2434
DB 2296 ATCTGACAAAGAGGCGCAGCGCG-CCCTGGAGCTCTGGGACCTGCTCCACTTCTCCA 2352
QY 2435 ATCAAGTTGCAAGGAATTTCTGAAATTTAAAGTCGTGTGTTTCAAGAGACCTGG 2494
DB 2353 GCCAAGTGGCTCAGGGCATGGCTTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2412
QY 2495 CCGCCAGGAAGCTGTGTTGTCACCCACCGGAAAGTGGTGAAGATATGTGACTTTTGGATTGG 2554
DB 2413 CAGCTCGAAACGCTGCTGTTGACCGAGCGGACATGTGGCCAAAGATTGGGAGCTTTGGAGCTGG 2472
QY 2555 CTGAGATATCATGAGTGATTCCAACTATGTGTGTCAGGGCAATGCCGCTCTGCTGTAA 2614
DB 2473 CTAGGACATCATGATGACTCCAACTATGTGTGTCAGGGCAATGCCGCTCTGCTGTAA 2532
QY 2615 AATGGATGGCCCCGAAAGCTGTTTGAAGGCATCTACACCATTTAAGAGTGAATGCTGCT 2674
DB 2533 AGTGGATGGCCCCCAGAGAGCATCTTTGACTGCTGTACACAGATTCAGAGTATGCTGCT 2592
QY 2675 CATATGGAATATTACTGTGGGAAATCTTCTCCTTGTGTGTAATCTTACCTGGCATTC 2734
DB 2593 CCTACGGCATCTCTCTCTGGAGATCTTCTCGTGTGTTCTGAAACCCCTACCCCGCATCC 2652
QY 2735 CGTTCGATGCTAACTTCTACAAACTGATTCAAAATGGATTAAATGATCAGCCATTTT 2794
DB 2653 TAGTGAACAAAGTTCTCAAACTGTTGTAAGGATGATACCAATGGCCAGCCTGTAT 2712
QY 2795 ATGCTACAGAAGAAATATACATTATATATGCAATCTCTGGGCTTTTGTGACTCAAGGAAAC 2854
DB 2713 TTGCACCGAAGACATATACAGCATCATGCAGTCTCTGCTGGAGCTTGAGCCTACCAGAA 2772
QY 2855 GGCCATCTCTCC 2866
DB 2773 GACCCACCTTCC 2784

Db 1681 ATCTATTCTATGCAACAATTGGTGTGTTGTCCTCTCTTCAATGTCGTTTTAAACCTTGCTA 1740
Qy 1741 ATTTGTACAAAGTACAAAAGCAATTTAGTATGAAAGCCAGCTACAGATGGTACAGGTG 1800
Db 1741 ATTTGTACAAAGTACAAAAGCAATTTAGTATGAAAGCCAGCTACAGATGGTACAGGTG 1800
Qy 1801 ACCGGCTCTCAGATAATAGTACTTCTACGTTGATTTTACAGAGAATATGAATATGATCTC 1860
Db 1801 ACCGGCTCTCAGATAATAGTACTTCTACGTTGATTTTACAGAGAATATGAATATGATCTC 1860
Qy 1861 AAATGGAGTTTCCAAAGAGAAAATTTAGAGTTTGGGAAGTACTAGGATCAGGTGCTTTTT 1920
Db 1861 AAATGGAGTTTCCAAAGAGAAAATTTAGAGTTTGGGAAGTACTAGGATCAGGTGCTTTTT 1920
Qy 1921 GGAAGTATGATGAACCAAGCAGCTTATGGAATTTAGCAAAACAGAGATCTCAATCCAGTTT 1980
Db 1921 GGAAGTATGATGAACCAAGCAGCTTATGGAATTTAGCAAAACAGAGATCTCAATCCAGTTT 1980
Qy 1981 GCGCTCAAAATGCTGAAGAAAAGCAGACAGCTCTGAAAGAGAGCACTCATGTTCAGAA 2040
Db 1981 GCGCTCAAAATGCTGAAGAAAAGCAGACAGCTCTGAAAGAGAGCACTCATGTTCAGAA 2040
Qy 2041 CTCAAGATGATGACCCAGCTGGGAAGCCAGAGAAATTTGTGAACCTGCTGGGGGGTGC 2100
Db 2041 CTCAAGATGATGACCCAGCTGGGAAGCCAGAGAAATTTGTGAACCTGCTGGGGGGTGC 2100
Qy 2101 ACACTGTACGACCAATTTACTTTGATTTTGAATCTGTTGATCTGTTGATCTTCAAC 2160
Db 2101 ACACTGTACGACCAATTTACTTTGATTTTGAATCTGTTGATCTGTTGATCTTCAAC 2160
Qy 2161 TATCTAAGAAATGAAAGAGAAAATTTACAGAGCTTGGACAGAGATTTTCAAGGAACAC 2220
Db 2161 TATCTAAGAAATGAAAGAGAAAATTTACAGAGCTTGGACAGAGATTTTCAAGGAACAC 2220
Qy 2221 AATTTTCAGTTTACCCCACTTCCAAATCCAAATCCAAATCCAAATCCAAATCCAAATCC 2280
Db 2221 AATTTTCAGTTTACCCCACTTCCAAATCCAAATCCAAATCCAAATCCAAATCCAAATCC 2280
Qy 2281 GAAGTTTCAGATACACCCGAGCTCGGATCAAATCTCAGGGCTTCATGGGAATTCATTTTAC 2340
Db 2281 GAAGTTTCAGATACACCCGAGCTCGGATCAAATCTCAGGGCTTCATGGGAATTCATTTTAC 2340
Qy 2341 TCTGAAGATGAATTTGAATATGAATAAACCAGAGCTGGAGAGAGAGAGAGAGAGAGAGAG 2400
Db 2341 TCTGAAGATGAATTTGAATATGAATAAACCAGAGCTGGAGAGAGAGAGAGAGAGAGAGAG 2400
Qy 2401 GTGCTTTACATTTGAAGATCTTCTTCTGTTGATATCAAGTTGCCAAAGGATGGAATTT 2460
Db 2401 GTGCTTTACATTTGAAGATCTTCTTCTGTTGATATCAAGTTGCCAAAGGATGGAATTT 2460
Qy 2461 CTGGAATTTAAGTCGTGTGTTTACAGAGACCTGGCCGCCAGGAACGTTGTTGCCACCCAC 2520
Db 2461 CTGGAATTTAAGTCGTGTGTTTACAGAGACCTGGCCGCCAGGAACGTTGTTGCCACCCAC 2520
Qy 2521 GGGAAAGTGGTGAAGATATGATGATTTGGATTTGGCTCGAGATATCATAGTATGATTTCAAC 2580
Db 2521 GGGAAAGTGGTGAAGATATGATGATTTGGATTTGGCTCGAGATATCATAGTATGATTTCAAC 2580
Qy 2581 TATGTTGTACGGGGCAATGCGCTGCTGTAATGATGATGATGATGATGATGATGATGATGAT 2640
Db 2581 TATGTTGTACGGGGCAATGCGCTGCTGTAATGATGATGATGATGATGATGATGATGATGAT 2640
Qy 2641 GAAGGCATCTACACCAATTAAGAGTATGATGCTGATGATGATGATGATGATGATGATGATG 2700
Db 2641 GAAGGCATCTACACCAATTAAGAGTATGATGCTGATGATGATGATGATGATGATGATGATG 2700
Qy 2701 TTCTCACTTGGTGAATCTTACCTGGCATTCGGTTGATGCTTAATCTTACAAACTG 2760
Db 2701 TTCTCACTTGGTGAATCTTACCTGGCATTCGGTTGATGCTTAATCTTACAAACTG 2760
Qy 2761 ATTCAAAATGGATTTAAATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2820
Db 2761 ATTCAAAATGGATTTAAATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2820

Qy 2821 ATGCAATCTGCTGGGCTTTTGAATCAAGAAAGCGCAATCTTCCCTTAATTTGACTTCG 2880
Db 2821 ATGCAATCTGCTGGGCTTTTGAATCAAGAAAGCGCAATCTTCCCTTAATTTGACTTCG 2880
Qy 2881 TTTTATGAGATGTCAAGCTGGCAGATGCAAGAAAGCGATGATCAGAATGTGATGCGCCGT 2940
Db 2881 TTTTATGAGATGTCAAGCTGGCAGATGCAAGAAAGCGATGATCAGAATGTGATGCGCCGT 2940
Qy 2941 GTTTCGGAATGTCTCTCACTACCAAAACAGGCGACCTTTTCCAGCAGAGATGATTTG 3000
Db 2941 GTTTCGGAATGTCTCTCACTACCAAAACAGGCGACCTTTTCCAGCAGAGATGATTTG 3000
Qy 3001 GGGCTACTCTCTCGCAGGCTCAGGTCGAAGATTCGTAGAGGAACAATTTAGTTTAAAG 3060
Db 3001 GGGCTACTCTCTCGCAGGCTCAGGTCGAAGATTCGTAGAGGAACAATTTAGTTTAAAG 3060
Qy 3061 ACTTCATCCCTCCACCTATCCCTAAACAGGCTGTAGATTACCAAAACAGATTAATTTTCA 3120
Db 3061 ACTTCATCCCTCCACCTATCCCTAAACAGGCTGTAGATTACCAAAACAGATTAATTTTCA 3120
Qy 3121 CACTAAAAGAAAATCTATTATCAACTGCTCTTCCAGCAGACTTTTCTCTAGAAGCCGTCT 3180
Db 3121 CACTAAAAGAAAATCTATTATCAACTGCTCTTCCAGCAGACTTTTCTCTAGAAGCCGTCT 3180
Qy 3181 GCGTTTACTCTTCTTTTCAAAAGGAGCTTTTGTAAATCAAATCATCTCTCACAAGGCAG 3240
Db 3181 GCGTTTACTCTTCTTTTCAAAAGGAGCTTTTGTAAATCAAATCATCTCTCACAAGGCAG 3240
Qy 3241 GAGGAGCTGATATGAATCTTTTATGGAGCATTTGATCTGCAATCCAGGCTTCTCAGGCG 3300
Db 3241 GAGGAGCTGATATGAATCTTTTATGGAGCATTTGATCTGCAATCCAGGCTTCTCAGGCG 3300
Qy 3301 GCTTGAGTGAATGTCTGCTGAGTACAGTATATTTTCTGTAATACATAAAACAAAGC 3360
Db 3301 GCTTGAGTGAATGTCTGCTGAGTACAGTATATTTTCTGTAATACATAAAACAAAGC 3360
Qy 3361 ATTTTCTTAAGAGAGCTAAATGATTTTTTAAAGTCTATGTTTAAATAATATGTAAA 3420
Db 3361 ATTTTCTTAAGAGAGCTAAATGATTTTTTAAAGTCTATGTTTAAATAATATGTAAA 3420
Qy 3421 TTTTTCAGCTATTTAGTATATATTTTATGGGTGGGAATAAAATTTCTACTACAGAAAA 3480
Db 3421 TTTTTCAGCTATTTAGTATATATTTTATGGGTGGGAATAAAATTTCTACTACAGAAAA 3480
Qy 3481 AAAAAAAGAAAAAAGAAAAA 3501
Db 3481 AAAAAAAGAAAAAAGAAAAA 3501

RESULT 2

AAQ79069

ID AAQ79069 standard; cDNA; 3501 BP.

XX

AC AAQ79069;

XX

DT 25-MAR-2003 (updated)

DT

04-JUL-1995 (first entry)

XX

DE Human flk-2 cDNA.

XX

KW Fetal liver kinase-2; flk-2; protein tyrosine-kinase receptor;

XX

hematopoiesis; stem cell; ds.

XX

OS Homo sapiens.

XX

FH Key Location/Qualifiers

FT CDS

FT /tag= a

FT sig_peptide

FT 58..138

FT /tag= b

FT mat_peptide

FT 139..3036

FT /tag= c

KW differentiation; treatment; anaemia; bone marrow damage;
 KW cancer chemotherapy; radiation; ds.
 XX Mus musculus.

OS
 PH Key Location/Qualifiers
 FT CDS 58..3039 /*tag= a
 FT sig_peptide 58..138 /*tag= b
 FT mat_peptide 139..3036 /*tag= c
 FT
 XX
 PN US5548065-A.
 XX
 XX
 PD 20-AUG-1996.
 XX
 PF 31-OCT-1994; 94US-0252517.
 XX
 PR 19-NOV-1992; 92US-0977451.
 PR 02-APR-1991; 91US-0679666.
 PR 28-JUN-1991; 91US-0728913.
 PR 15-NOV-1991; 91US-0793065.
 PR 24-DEC-1991; 91US-0813593.
 PR 26-JUN-1992; 92US-0906397.
 PR 12-NOV-1992; 92US-0975049.
 PR 30-APR-1993; 93US-0055269.
 PR 31-OCT-1994; 94US-0252517.
 XX
 PA (UYPR-) UNIV PRINCETON.
 XX
 XX Lemischka IR;
 PI
 XX WPI; 1996-392678/39.
 XX P-PSDB; AAR97419.
 DR
 DR
 XX

Anti-fetal liver kinase 2 (flk-2) antibodies - useful in assays,
 for isolating haematopoietic stem cells expressing receptor and for
 obtaining ligands

Disclosure; Columns 39-48; 50pp; English.

CC The present sequence encodes murine fetal liver kinase 2 (flk-2),
 CC a protein tyrosine kinase. Isolated antibodies, pref. monoclonal,
 CC raised against the extracellular portion of flk-2 can be used to
 CC assay for flk receptors on the surface of primitive haematopoietic
 CC stem cells, and to isolate positive cells. The antibodies can also
 CC be used as, or to obtain ligands, which stimulate the proliferation
 CC and/or differentiation of stem cells. The ligands can be used, e.g.
 CC for treating anaemia, or bone marrow damage resulting from cancer
 CC chemotherapy, or radiation.
 CC (Updated on 25-MAR-2003 to correct PF field.)
 XX

XX SQ Sequence 3501 BP; 1068 A; 709 C; 784 G; 940 T; 0 other;

Query Match 100.0%; Score 3501; DB 17; Length 3501;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 3501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGAGCGGCATCCGAGGCTGGCGCGCCCTGGGGACCCCGGGCTCCGGAGCCATG 60
 Db 1 CGAGCGGCATCCGAGGCTGGCGCGCCCTGGGGACCCCGGGCTCCGGAGCCATG 60
 Qy 61 CCGCGTTGGCGCGGACCGGACCCCTGGCGTCTGTTTTCGCAATGATA 120
 Db 61 CCGCGTTGGCGCGGACCGGACCCCTGGCGTCTGTTTTCGCAATGATA 120
 Qy 121 TTTCGGACTATTACAAATCAAGATCTGCTGATCAAGTGTGTTTAAATCAATCAAG 180
 Db 121 TTTCGGACTATTACAAATCAAGATCTGCTGATCAAGTGTGTTTAAATCAATCAAG 180
 Qy 181 AACCAATGATTCATAGTGGGAAGTCATCATATCCCATGATCAGAAATCCCGGAA 240
 Db 181 AACCAATGATTCATAGTGGGAAGTCATCATATCCCATGATCAGAAATCCCGGAA 240

Db 181 AACCAATGATTCATAGTGGGAAGTCATCATATCCCATGATCAGAAATCCCGGAA 240
 Qy 241 GACCTCGGGTGTGCGTTGAGACCCAGAGCTCAGGACAGTGTACGAAAGCTCCCGCTGTG 300
 Db 241 GACCTCGGGTGTGCGTTGAGACCCAGAGCTCAGGACAGTGTACGAAAGCTCCCGCTGTG 300
 Qy 301 GAAGTGGATGATCTGCTTCCATCAGTCAAGTGTGCTGATGCCCCAGGGAACATT 360
 Db 301 GAAGTGGATGATCTGCTTCCATCAGTCAAGTGTGCTGATGCCCCAGGGAACATT 360
 Qy 361 TCCTGTCTCTGGGTCTTTAAGCAGAGCTCCCTGANTGCCAGCCACATTTTCAATTACAA 420
 Db 361 TCCTGTCTCTGGGTCTTTAAGCAGAGCTCCCTGANTGCCAGCCACATTTTCAATTACAA 420
 Qy 421 AACAGAGAGTGTGTTTCCATGTCATTTGAAATGACAGAAACCAAGCTCGAGAAATAC 480
 Db 421 AACAGAGAGTGTGTTTCCATGTCATTTGAAATGACAGAAACCAAGCTCGAGAAATAC 480
 Qy 481 CTACTTTTATTCAGAGTGAAGCTACCAATACCAATATTTGTTTACAGTGAATTAAGA 540
 Db 481 CTACTTTTATTCAGAGTGAAGCTACCAATATTTGTTTACAGTGAATTAAGA 540
 Qy 541 AATACCCCTGCTTTACACATTAAGAAGACCTTTACTTTAGAAAATGGAACACAGGACGCC 600
 Db 541 AATACCCCTGCTTTACACATTAAGAAGACCTTTACTTTAGAAAATGGAACACAGGACGCC 600
 Qy 601 CTGCTCTGCATATCTGAGAGCGTTCCAGAGCGGATCGTGGAAATGGTGTCTTTGCCATTCA 660
 Db 601 CTGCTCTGCATATCTGAGAGCGTTCCAGAGCGGATCGTGGAAATGGTGTCTTTGCCATTCA 660
 Qy 661 CAGGGGAAAGCTGTAAAGAAAGTCCAGCTGTTGTTTAAAGAGGAGGAAAAAGTGCTT 720
 Db 661 CAGGGGAAAGCTGTAAAGAAAGTCCAGCTGTTGTTTAAAGAGGAGGAAAAAGTGCTT 720
 Qy 721 CATGAATTTTGGGACGACATAAGTGTCTGTCGACAGAAATGAATGGGACGGAATGC 780
 Db 721 CATGAATTTTGGGACGACATAAGTGTCTGTCGACAGAAATGAATGGGACGGAATGC 780
 Qy 781 ACCAGGCTGTTTCAAAATAGATCTAAATCAAACTCCTCAGACCACATTTGCCAAATTTT 840
 Db 781 ACCAGGCTGTTTCAAAATAGATCTAAATCAAACTCCTCAGACCACATTTGCCAAATTTT 840
 Qy 841 CTTAAAGTAGGGGAACCCCTTATGGATTAAGTGCAGAAAGCTGTTTCTGTAACCATGATTC 900
 Db 841 CTTAAAGTAGGGGAACCCCTTATGGATTAAGTGCAGAAAGCTGTTTCTGTAACCATGATTC 900
 Qy 901 GGGCTCAGCTGGGAATTAAGAAAACAGACATCGAGAGGGGCAACTACTTTGAGATGAGT 960
 Db 901 GGGCTCAGCTGGGAATTAAGAAAACAGACATCGAGAGGGGCAACTACTTTGAGATGAGT 960
 Qy 961 ACCTATTCAACAAACAGAACTATGATACGGATCTGTTTGTGTTTGTATCATCAGTGGCA 1020
 Db 961 ACCTATTCAACAAACAGAACTATGATACGGATCTGTTTGTGTTTGTATCATCAGTGGCA 1020
 Qy 1021 AGAAACGACACCGGATCTACACTTGTTCCTCTTCAAGAGCATCCAGTCAATCAGCTTTG 1080
 Db 1021 AGAAACGACACCGGATCTACACTTGTTCCTCTTCAAGAGCATCCAGTCAATCAGCTTTG 1080
 Qy 1081 GTTACCATCTGTAGGAAGGGATTTATAAATGCTACCAATTCAGTGAAGATTAATGAAATT 1140
 Db 1081 GTTACCATCTGTAGGAAGGGATTTATAAATGCTACCAATTCAGTGAAGATTAATGAAATT 1140
 Qy 1141 GACCAATATGAAGTGTGTTTCTGTCAGTGTAAAGCTCCCAACAAATCAGATGT 1200
 Db 1141 GACCAATATGAAGTGTGTTTCTGTCAGTGTAAAGCTCCCAACAAATCAGATGT 1200
 Qy 1201 ACCTGGAGCTTCTCTCGAAAATCAATTTCTTGTGAGCAAAAAGGGTCTTGATAACGGATAC 1260
 Db 1201 ACCTGGAGCTTCTCTCGAAAATCAATTTCTTGTGAGCAAAAAGGGTCTTGATAACGGATAC 1260
 Qy 1261 AGCATATCCAAAGTTTTCGAATCATAAGCACAGCCAGGAGGAATATATATTCATCCAGAA 1320
 Db 1261 AGCATATCCAAAGTTTTCGAATCATAAGCACAGCCAGGAGGAATATATATTCATCCAGAA 1320


```

Db      3481 AAAAAAAAAAAAAAAAAAAAAA 3501
RESULT 4
AAT72117
ID      AAT72117 standard; cDNA; 3501 BP.
XX      AC
XX      AAT72117;
XX
XX      25-MAR-2003 (updated)
XX      19-AUG-1997 (first entry)
XX
XX      Human flk-2 receptor coding sequence.
XX
XX      Human; fetal liver kinase 2; flk2; receptor.protein tyrosine kinase;
KW      PKT; liver; spleen; thymus; adult; brain; bone marrow; primitive portion;
KW      haematopoietic hierarchy; extracellular domain; soluble form; ligand;
KW      proliferation; differentiation; mammalian; haematopoietic stem cell;
KW      macrocytic anaemia; aplastic anaemia; cancer; radiation; mouse; ss.
XX
XX      Homo sapiens.
XX
XX      Key Location/Qualifiers
XX      CDS 58..3039
XX      /tag= a
XX      /product= Human_flk-2_receptor
XX      sig_peptide 58..138
XX      /tag= b
XX      mat_peptide 139..3036
XX      /tag= c
XX
XX      US5621090-A.
XX
XX      15-APR-1997.
XX
XX      26-JUN-1992; 92US-0906397.
XX
XX      26-JUN-1992; 92US-0906397.
XX      02-APR-1991; 91US-0679666.
XX      28-JUN-1991; 91US-0728913.
XX      15-NOV-1991; 91US-0793065.
XX      24-DEC-1991; 91US-0813593.
XX
XX      (UYPR-) UNIV PRINCETON.
XX
XX      Lemischka IR;
XX
XX      WPI; 1997-235228/21.
XX      P-PSDB; AAW19873.
XX
XX      Protein containing the extracellular domain of human flk-2 - used
XX      for identification of primitive haematopoietic cell proliferation
XX      and differentiation stimulatory ligands, e.g. for treating anaemia
XX
XX      Claim 1; Fig 1B; 55pp; English.
XX
XX      This sequence encodes the human fetal liver kinase 2 (flk2). flk-2 is
XX      a receptor protein tyrosine kinase (PTK) and is important in transducing
XX      putative self-renewal signals from the environment. flk-2 is expressed
XX      in fetal liver, spleen, fetal thymus, adult brain and adult bone marrow,
XX      and it is thought that flk-2 is expressed in the entire primitive portion
XX      of the haematopoietic hierarchy. The invention concerns a recombinant
XX      nucleic acid, preferably mRNA, which encodes a protein containing only
XX      the extracellular domain of human flk-2 and lacking the flk-2 intra-
XX      cellular catalytic domain. The resultant protein represents a soluble
XX      form of flk-2 which is used to isolate specific ligands for flk-2. These
XX      ligands can be used to stimulate proliferation and/or differentiation of
XX      mammalian haematopoietic stem cells, in vivo or in vitro, e.g. for
XX      treatment of macrocytic or aplastic anaemia or bone marrow damage caused
XX      by cancer treatment or radiation.
XX      (Updated on 25-MAR-2003 to correct PF field.)
XX

```

```

SQ      Sequence 3501 BP; 1068 A; 709 C; 784 G; 940 T; 0 other;
Query Match      100.0%; Score 3501; DB 18; Length 3501;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      1 CGAGCGGCATCCGAGGGCTGGGCGCCCTGGGGACCCCGGGCTCCGGAGGCCATG 60
DB      1 CGAGCGGCATCCGAGGGCTGGGCGCCCTGGGGACCCCGGGCTCCGGAGGCCATG 60
QY      61 CCGGGTGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 120
DB      61 CCGGGTGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 120
QY      121 TTTGGGACTATTACAAATCAAGATCTGCTGTGATCAAGTGTGTTTAAATCAATCAATAG 180
DB      121 TTTGGGACTATTACAAATCAAGATCTGCTGTGATCAAGTGTGTTTAAATCAATCAATAG 180
QY      181 AACAAATGATTCATCAGTGGGGAAGTCAATCATATATCCATGATATCAAGATCCCGGAA 240
DB      181 AACAAATGATTCATCAGTGGGGAAGTCAATCATATATCCATGATATCAAGATCCCGGAA 240
QY      241 GACCTCGGGTGTGGTGGAGACCCCGGAGCTCAGGACAGTGTACGAGCTGCCGCTGTG 300
DB      241 GACCTCGGGTGTGGTGGAGACCCCGGAGCTCAGGACAGTGTACGAGCTGCCGCTGTG 300
QY      301 GAAGTGATGTATCTGCTTCCATCACACTGCAAGTGTGTCGATGCCCGGGAACATT 360
DB      301 GAAGTGATGTATCTGCTTCCATCACACTGCAAGTGTGTCGATGCCCGGGAACATT 360
QY      361 TCCTGTCTCTGGGTCTTTAAGCAGAGCTCCCTGAATTCGCCAGCCACATTTGATTTACAA 420
DB      361 TCCTGTCTCTGGGTCTTTAAGCAGAGCTCCCTGAATTCGCCAGCCACATTTGATTTACAA 420
QY      421 AACAGAGGAGTGTTCCTGTCATTTTGAATGACAGAAACCCAGACTGGGAATAC 480
DB      421 AACAGAGGAGTGTTCCTGTCATTTTGAATGACAGAAACCCAGACTGGGAATAC 480
QY      481 CTACTTTTATTTCAGAGTGAAGCTACCAATTTACAAATTTTGTTCAGTGTAGTATAAGA 540
DB      481 CTACTTTTATTTCAGAGTGAAGCTACCAATTTACAAATTTTGTTCAGTGTAGTATAAGA 540
QY      541 AATACCTCTGTTTACATTAAGAACGCTTACTTTAGAAAAATGAAAAACCGAGCGCC 600
DB      541 AATACCTCTGTTTACATTAAGAACGCTTACTTTAGAAAAATGAAAAACCGAGCGCC 600
QY      601 CTGCTCTGCATATCTGAGAGCGTTCCAGAGCGGATCGTGGATGGGTGCTTGCATTTCA 660
DB      601 CTGCTCTGCATATCTGAGAGCGTTCCAGAGCGGATCGTGGATGGGTGCTTGCATTTCA 660
QY      661 CAGGGGAAAGCTGTAAAGAAAGTCCAGCTGTTGTTTAAAAAGGAGGAAAAAGTGCTT 720
DB      661 CAGGGGAAAGCTGTAAAGAAAGTCCAGCTGTTGTTTAAAAAGGAGGAAAAAGTGCTT 720
QY      721 CATGAATTTATTGGGACCGGACATTAAGGTGTGTGCCAGAAATGAATCGGGGAGGAATGC 780
DB      721 CATGAATTTATTGGGACCGGACATTAAGGTGTGTGCCAGAAATGAATCGGGGAGGAATGC 780
QY      781 ACCAGGCTGTTTACAAATAGATCTAAATCAAACTCTCCAGACCAATTTGCCCAATTTT 840
DB      781 ACCAGGCTGTTTACAAATAGATCTAAATCAAACTCTCCAGACCAATTTGCCCAATTTT 840
QY      841 CTTAAAGTGGGGAACCCCTATGATAGGTGCAAGCTGTTTCATGTCAACCATGATTC 900
DB      841 CTTAAAGTGGGGAACCCCTATGATAGGTGCAAGCTGTTTCATGTCAACCATGATTC 900
QY      901 GGGCTCACCTGGGAATTTAGAAAAACAAAGCACTCGAGGAGGCAACTACTTTGAGATGAGT 960
DB      901 GGGCTCACCTGGGAATTTAGAAAAACAAAGCACTCGAGGAGGCAACTACTTTGAGATGAGT 960
QY      961 ACCTATTCAACAAACAGAACTATGATACGGATTCCTGTTTGTGTTTGTATCATCATGCGCA 1020
DB      961 ACCTATTCAACAAACAGAACTATGATACGGATTCCTGTTTGTGTTTGTATCATCATGCGCA 1020

```

1021 AGAAACGACACCGGATACACTGTTCTCTTCAAAGCATCCAGTCAATCAGCTTTG 1080
1021 AGAAACGACACCGGATACACTGTTCTCTTCAAAGCATCCAGTCAATCAGCTTTG 1080
1081 GTTACCAATCGTAGGAAAGGGATTTATAAATGCTACCAATTCAGGTGAAGATTGAAAT 1140
1081 GTTACCAATCGTAGGAAAGGGATTTATAAATGCTACCAATTCAGGTGAAGATTGAAAT 1140
1141 GACCAATATGAGAGATTTGTTTCTCTGTCAGGTTTAAAGCTTACCCACAATCGATGT 1200
1141 GACCAATATGAGAGATTTGTTTCTCTGTCAGGTTTAAAGCTTACCCACAATCGATGT 1200
1201 ACCTGGACCTTCTCTCGAAATCAATTTCTGTCAGCAAAAAGGTCTTGATAACGGATAC 1260
1201 ACCTGGACCTTCTCTCGAAATCAATTTCTGTCAGCAAAAAGGTCTTGATAACGGATAC 1260
1261 AGCATATCCAAAGTTTGCATATCAPAAGCACACGACCGAGAGATATATATTCATCGAGAA 1320
1261 AGCATATCCAAAGTTTGCATATCAPAAGCACACGACCGAGAGATATATATTCATCGAGAA 1320
1321 AATGATGATGCCCAATTTACCAAAATGTTACGCTGAAATATAAGAGAAACCTCAAGTG 1380
1321 AATGATGATGCCCAATTTACCAAAATGTTACGCTGAAATATAAGAGAAACCTCAAGTG 1380
1381 CTGCGAAGCATCGGCAAGTFCAGCGCTCTGTTTCTCGGATGATACCCATACCATCT 1440
1381 CTGCGAAGCATCGGCAAGTFCAGCGCTCTGTTTCTCGGATGATACCCATACCATCT 1440
1441 TGGACCTGGAAGAGTGTTCAGACAAAGTCTCCCAATGCGACAGAGAGATCACAGAAGGA 1500
1441 TGGACCTGGAAGAGTGTTCAGACAAAGTCTCCCAATGCGACAGAGAGATCACAGAAGGA 1500
1501 GTCTGGAATAGAAGGCTAACAGAAAAGTGTGGAAGTGTGCGAGTACTCTA 1560
1501 GTCTGGAATAGAAGGCTAACAGAAAAGTGTGGAAGTGTGCGAGTACTCTA 1560
1561 AACATGAGTGAAGCCATAAAGGGTTCCTGTCAGTCTGTCGATACAAATCCCTTGGC 1620
1561 AACATGAGTGAAGCCATAAAGGGTTCCTGTCAGTCTGTCGATACAAATCCCTTGGC 1620
1621 ACATCTTGTGAGACGATCTCTTAAACTCTCCAGGCCCTTCCCTTTCATCCAGACAAC 1680
1621 ACATCTTGTGAGACGATCTCTTAAACTCTCCAGGCCCTTCCCTTTCATCCAGACAAC 1680
1681 ATCTCAATCTATGCAACAATTTGGTGTGTCCTCTCTTCAATTTGCTGTTTAAACCTGCTA 1740
1681 ATCTCAATCTATGCAACAATTTGGTGTGTCCTCTCTTCAATTTGCTGTTTAAACCTGCTA 1740
1741 ATTTGTCAAGTACAAAAGCAATTTAGGTATGAAGCCAGCTACAGATGGTACAGGTG 1800
1741 ATTTGTCAAGTACAAAAGCAATTTAGGTATGAAGCCAGCTACAGATGGTACAGGTG 1800
1801 ACCGGCTCTCAGATAATGAGTACTTCTACGTTGATTTACAGAGATATGAATGATCTC 1860
1801 ACCGGCTCTCAGATAATGAGTACTTCTACGTTGATTTACAGAGATATGAATGATCTC 1860
1861 AAATGGAGTTTCCAAAGAGAAATTTAGAGTTTGGAGGTACTAGGATCAGGTGCTTTT 1920
1861 AAATGGAGTTTCCAAAGAGAAATTTAGAGTTTGGAGGTACTAGGATCAGGTGCTTTT 1920
1921 GGAAGTATGATGACCAAGCTTATGNAATTTAGCAAAACAGGAGTCTCAATCCAGGTT 1980
1921 GGAAGTATGATGACCAAGCTTATGNAATTTAGCAAAACAGGAGTCTCAATCCAGGTT 1980
1981 GCCGTCAAAATGCTGAAAGAAAGCAGACAGCTCTGGAAGAGAGCACTCATGTGAGAA 2040
1981 GCCGTCAAAATGCTGAAAGAAAGCAGACAGCTCTGGAAGAGAGCACTCATGTGAGAA 2040
2041 CTCAAGATGATGACCCAGCTGGGAAGCCACAGAAATTTGTGAACCTGTGGGGGGTGC 2100
2041 CTCAAGATGATGACCCAGCTGGGAAGCCACAGAAATTTGTGAACCTGTGGGGGGTGC 2100

2101 ACCTGTGAGGACCAATTTACTTGAATTTTGAATACTGTGCTATGGTGATCTTCTCAAC 2160
2101 ACCTGTGAGGACCAATTTACTTGAATTTTGAATACTGTGCTATGGTGATCTTCTCAAC 2160
2161 TATCTAAGAAAGTAAAGAGAAAAATTTTCAAGGACTTGGACAGAGATTTTCAAGGAACAC 2220
2161 TATCTAAGAAAGTAAAGAGAAAAATTTTCAAGGACTTGGACAGAGATTTTCAAGGAACAC 2220
2221 AATTTCAAGTTTATACCCACCTTTCATATCAATCCAAATTCAGCATGCTGCTTCAAGA 2280
2221 AATTTCAAGTTTATACCCACCTTTCATATCAATCCAAATTCAGCATGCTGCTTCAAGA 2280
2281 GAAGTTCAGATACACCCGACCTCGGATCAAAATCTCAGGGCTTCATGGGAAATTCATTTCA 2340
2281 GAAGTTCAGATACACCCGACCTCGGATCAAAATCTCAGGGCTTCATGGGAAATTCATTTCA 2340
2341 TCTGAAGATGAAATTTGAATATGAAAAACAAAAAGGCTGGAAAGAGAGAGGACTTGAAT 2400
2341 TCTGAAGATGAAATTTGAATATGAAAAACAAAAAGGCTGGAAAGAGAGAGGACTTGAAT 2400
2401 GTGCTTACATTTGAAGATCTTCTTGTGCTTTCATATCAAGTTGCCAAAGGAATGGAATTT 2460
2401 GTGCTTACATTTGAAGATCTTCTTGTGCTTTCATATCAAGTTGCCAAAGGAATGGAATTT 2460
2461 CTGGAATTTAAGTCTGCTGTTTACAGAGACCTGCGCCAGAGAGCTGCTTGTCAACCCAC 2520
2461 CTGGAATTTAAGTCTGCTGTTTACAGAGACCTGCGCCAGAGAGCTGCTTGTCAACCCAC 2520
2521 GGGAAAGTGTGGAAGATATGTGACTTTTGGATGCTCGAGATATCATGAGTGAATTTCCAAC 2580
2521 GGGAAAGTGTGGAAGATATGTGACTTTTGGATGCTCGAGATATCATGAGTGAATTTCCAAC 2580
2581 TATGTTTCTCAGGGGCAATGCCCCTGCTGCTGTAATAATGGATGGCCCCCGAAAGCCCTGTTT 2640
2581 TATGTTTCTCAGGGGCAATGCCCCTGCTGCTGTAATAATGGATGGCCCCCGAAAGCCCTGTTT 2640
2641 GAAGGCATCTACACCATTAAGAGTGTCTGTCATATGGAATATTAATCTGTGGAAATC 2700
2641 GAAGGCATCTACACCATTAAGAGTGTCTGTCATATGGAATATTAATCTGTGGAAATC 2700
2701 TTCTCACTGCTGTGAAATCTTACCTCCCTGGATTCGGGTGATGCTTAACACTTCAAGACTG 2760
2701 TTCTCACTGCTGTGAAATCTTACCTCCCTGGATTCGGGTGATGCTTAACACTTCAAGACTG 2760
2761 ATTCAAAATGGAATTTAAATGGATCAGCCATTTTATGCTACAGAGAAATATACATATATA 2820
2761 ATTCAAAATGGAATTTAAATGGATCAGCCATTTTATGCTACAGAGAAATATACATATATA 2820
2821 ATGCAATCTCTGCTGGGCTTTTGAATCAAGGAAACGGCCATCTTCCCTAATTTGACTTCG 2880
2821 ATGCAATCTCTGCTGGGCTTTTGAATCAAGGAAACGGCCATCTTCCCTAATTTGACTTCG 2880
2881 TTTTATAGGATGTCAGCTGGGAGATGAGAGAGAGAGATGATCAAGATGTTGGATGGCGGT 2940
2881 TTTTATAGGATGTCAGCTGGGAGATGAGAGAGAGAGATGATCAAGATGTTGGATGGCGGT 2940
2941 GTTTCCGAATGTCCTCAACCTTACCAAAACAGCGACCTTTTACAGCAGAGAGATGGAATTTG 3000
2941 GTTTCCGAATGTCCTCAACCTTACCAAAACAGCGACCTTTTACAGCAGAGAGATGGAATTTG 3000
3001 GGGCTACTCTCTCCGAGGCTCAGGTGGAAGATTCGTAGAGGAACAATTTAGTTTTAAGG 3060
3001 GGGCTACTCTCTCCGAGGCTCAGGTGGAAGATTCGTAGAGGAACAATTTAGTTTTAAGG 3060
3061 ACTTCATCTCTCCACCTATCCCTTAAAGGCTGTAGATTACCAAAACAGAGTAAATTTTCA 3120
3061 ACTTCATCTCTCCACCTATCCCTTAAAGGCTGTAGATTACCAAAACAGAGTAAATTTTCA 3120
3121 CACTAAAGAAATCTATTATCACTGCTTCCAGCACTTTTCTCTAGAGAGCCCTCT 3180
3121 CACTAAAGAAATCTATTATCACTGCTTCCAGCACTTTTCTCTAGAGAGCCCTCT 3180
3181 GCGTTTACTCTTGTGTTTCAAAGGGGACTTTTGTAAAAATCAATCATCTCTGTCAAGGCAG 3240

Db 841 CTTAAAGTAGGGAAACCTTTATGGATAAGGTGCAAGCTGTTGATGTAACCATGGATTC 900
Qy 901 GGCGTCACTGGGAAATTAGAAAACAAGCACTCGAGGAGGCAACTACTTTGAGATGAGT 960
Db 901 GGCGTCACTGGGAAATTAGAAAACAAGCACTCGAGGAGGCAACTACTTTGAGATGAGT 960
Qy 961 ACCTATTCAACAACAGAACTATGATACGGATCTCTGTTGCTTTGATCATCAGTGGCA 1020
Db 961 ACCTATTCAACAACAGAACTATGATACGGATCTCTGTTGCTTTGATCATCAGTGGCA 1020
Qy 1021 AGAAACGACACCGGATACACTGTTCTCTTCAAAGCATCCAGTCAATCAGCTTTG 1080
Db 1021 AGAAACGACACCGGATACACTGTTCTCTTCAAAGCATCCAGTCAATCAGCTTTG 1080
Qy 1081 GTTACCATCGTAGAAGGATTTAATAATGCTACCAATTCAGTGGAAGATTATGAAT 1140
Db 1081 GTTACCATCGTAGAAGGATTTAATAATGCTACCAATTCAGTGGAAGATTATGAAT 1140
Qy 1141 GACCAATATGAAGAGTTTCTCGAAAATCATTTCTGTCAGGTTTAAAGCCTACCCACAATCAGATGT 1200
Db 1141 GACCAATATGAAGAGTTTCTCGAAAATCATTTCTGTCAGGTTTAAAGCCTACCCACAATCAGATGT 1200
Qy 1201 AGCTGGACCTTCTCTCGAAAATCATTTCTGTCAGGTTTAAAGCCTACCCACAATCAGATGT 1260
Db 1201 AGCTGGACCTTCTCTCGAAAATCATTTCTGTCAGGTTTAAAGCCTACCCACAATCAGATGT 1260
Qy 1261 AGCATATCCAAGTTTGGCAATATAGCACACCGCAGGAGAAATATATTTCCATGCGAA 1320
Db 1261 AGCATATCCAAGTTTGGCAATATAGCACACCGCAGGAGAAATATATTTCCATGCGAA 1320
Qy 1321 AATGATGATGCCAATTTACCAAAATGTTTCAAGTGAATTAAGAGGAAACCTCAAGTG 1380
Db 1321 AATGATGATGCCAATTTACCAAAATGTTTCAAGTGAATTAAGAGGAAACCTCAAGTG 1380
Qy 1381 CTCGCAAGAGATCGGCAAGTCAAGGCTGCTGTTCTCGATGGATACCCATACCATCT 1440
Db 1381 CTCGCAAGAGATCGGCAAGTCAAGGCTGCTGTTCTCGATGGATACCCATACCATCT 1440
Qy 1441 TGGACCTGGAAGAAGTGTTCAGACAAGTCTCCAACTGCAAGAAGAGATCAAGAAGGA 1500
Db 1441 TGGACCTGGAAGAAGTGTTCAGACAAGTCTCCAACTGCAAGAAGAGATCAAGAAGGA 1500
Qy 1501 GTCTGGAATAGAAAGCTTAACAGAAAAGTGTGGAAGTGTGAGCAGTGTGAGCAGTACTCTA 1560
Db 1501 GTCTGGAATAGAAAGCTTAACAGAAAAGTGTGGAAGTGTGAGCAGTGTGAGCAGTACTCTA 1560
Qy 1561 AACATGAGTGAAGCCATAAAGGTTTCTGGTCAAGTGTGTCATACAAATTCCTTGGC 1620
Db 1561 AACATGAGTGAAGCCATAAAGGTTTCTGGTCAAGTGTGTCATACAAATTCCTTGGC 1620
Qy 1621 ACATCTTGTGAGACGATCCCTTTTAAACTCTCCAGGCCCTTCCCTTTCATCCAAAGCAAC 1680
Db 1621 ACATCTTGTGAGACGATCCCTTTTAAACTCTCCAGGCCCTTCCCTTTCATCCAAAGCAAC 1680
Qy 1681 ATCTCAATTCATGCAACAATTTGTTGTTGTTCTCTTCAATGTCGTTTAAACCTTGCTA 1740
Db 1681 ATCTCAATTCATGCAACAATTTGTTGTTGTTCTCTTCAATGTCGTTTAAACCTTGCTA 1740
Qy 1741 ATTTGTGACAGTACAAAAGCAATTTAGGTATGAAGCCAGTACAGATGTTACAGGTG 1800
Db 1741 ATTTGTGACAGTACAAAAGCAATTTAGGTATGAAGCCAGTACAGATGTTACAGGTG 1800
Qy 1801 ACCGGTCTCTCAGATAATGAGTACTTCTAGCTTGAATTCAGAGAATATGAATGATCTC 1860
Db 1801 ACCGGTCTCTCAGATAATGAGTACTTCTAGCTTGAATTCAGAGAATATGAATGATCTC 1860
Qy 1861 AAATGGAGTTTCCAAGAAAATTTAGAGTTTGGAAAGTACTAGGATCAGTGCTTTT 1920
Db 1861 AAATGGAGTTTCCAAGAAAATTTAGAGTTTGGAAAGTACTAGGATCAGTGCTTTT 1920
Qy 1921 GAAAAGTGAATGACCAACAGCTTATGGAATTTAGCAAAACAGGAGTCTCAATCCAGTTT 1980

Db 1921 GAAAAGTGAATGACCAACAGCTTATGGAATTTAGCAAAACAGGAGTCTCAATCCAGTTT 1980
Qy 1981 GGCGTCAAAATGCTGAAAGAAAAGCAGACAGCTCTGAAAGAGAGGCACTCATGTFCAGAA 2040
Db 1981 GGCGTCAAAATGCTGAAAGAAAAGCAGACAGCTCTGAAAGAGAGGCACTCATGTFCAGAA 2040
Qy 2041 CTCAGATGATGACCCAGTGGGAAGCCAGAGAAATTTGTAACCTGCTGGGGCGTGC 2100
Db 2041 CTCAGATGATGACCCAGTGGGAAGCCAGAGAAATTTGTAACCTGCTGGGGCGTGC 2100
Qy 2101 ACATGTCAGGACCAATTTACTTGAATTTTGAATCTGTTGCTATGTTGATCTTCTCAAC 2160
Db 2101 ACATGTCAGGACCAATTTACTTGAATTTTGAATCTGTTGCTATGTTGATCTTCTCAAC 2160
Qy 2161 TATCTAAGAGTAAAAGAGAAAATTTCAAGGACTTTGGACAGAGATTTTCAAGGAAAC 2220
Db 2161 TATCTAAGAGTAAAAGAGAAAATTTCAAGGACTTTGGACAGAGATTTTCAAGGAAAC 2220
Qy 2221 AATTTCAAGTAAAAGAGAAAATTTCAAGGACTTTGGACAGAGATTTTCAAGGAAAC 2280
Db 2221 AATTTCAAGTAAAAGAGAAAATTTCAAGGACTTTGGACAGAGATTTTCAAGGAAAC 2280
Qy 2281 GAAGTTTCAGATACACCCGAGTCCGATCAAAATCTCAGGGCTTTCATGGAATTTCAATTCAC 2340
Db 2281 GAAGTTTCAGATACACCCGAGTCCGATCAAAATCTCAGGGCTTTCATGGAATTTCAATTCAC 2340
Qy 2341 TCTGAAGATGAAATTTGAATATGAAAACCAAAAAGGCTGGAAGAGAGGAGGACTTGAAT 2400
Db 2341 TCTGAAGATGAAATTTGAATATGAAAACCAAAAAGGCTGGAAGAGAGGAGGACTTGAAT 2400
Qy 2401 GTGCTTACATTTGAAGATCTTCTTGTGCTTTCATCAAGTTCGCAAGGAATGGAATTT 2460
Db 2401 GTGCTTACATTTGAAGATCTTCTTGTGCTTTCATCAAGTTCGCAAGGAATGGAATTT 2460
Qy 2461 CTGGAATTTAAGTCTGTTTTCAGAGACCTCGCGCCAGGAAACGCTGTTGTCAACCCAC 2520
Db 2461 CTGGAATTTAAGTCTGTTTTCAGAGACCTCGCGCCAGGAAACGCTGTTGTCAACCCAC 2520
Qy 2521 GGAAGAGTGGTGAAGATATGTGACTTTGGATTTGGCTCGAGATATCATGAGTATTCCAAC 2580
Db 2521 GGAAGAGTGGTGAAGATATGTGACTTTGGATTTGGCTCGAGATATCATGAGTATTCCAAC 2580
Qy 2581 TATGTTGTGAGGGCAATGCGCTCTGCTGTTAAAATGGATGGCCCCCGAAAGCTGTTT 2640
Db 2581 TATGTTGTGAGGGCAATGCGCTCTGCTGTTAAAATGGATGGCCCCCGAAAGCTGTTT 2640
Qy 2641 GAAGGATCTACACCATTAAGAGTGTGCTGCTCATATGGAATATTTACTGTGGGAAATC 2700
Db 2641 GAAGGATCTACACCATTAAGAGTGTGCTGCTCATATGGAATATTTACTGTGGGAAATC 2700
Qy 2701 TTCTCACTTGGTGTGAATCTTACCTGCGCATTCGCTGCTGATCTAACTTCTACAAACTG 2760
Db 2701 TTCTCACTTGGTGTGAATCTTACCTGCGCATTCGCTGCTGATCTAACTTCTACAAACTG 2760
Qy 2761 ATTCAAAATGGATTTAAAATGGATCAGCCATTTTATGCTACAGAAGAAATATACATTATA 2820
Db 2761 ATTCAAAATGGATTTAAAATGGATCAGCCATTTTATGCTACAGAAGAAATATACATTATA 2820
Qy 2821 ATGCAATCTCTGCTGGGCTTTGATCTCAAGAAAACGGCCATCTTCCCTAATTTGACTTCG 2880
Db 2821 ATGCAATCTCTGCTGGGCTTTTGAATCTCAAGAAAACGGCCATCTTCCCTAATTTGACTTCG 2880
Qy 2881 TTTTATGAGTGTGAGTGGCAGATGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2940
Db 2881 TTTTATGAGTGTGAGTGGCAGATGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2940
Qy 2941 GTTTCGGAATGTCTCTCACACCTTACCAAAACAGGCGACCTTTTTCAGCAGAGAGATGGATTTG 3000
Db 2941 GTTTCGGAATGTCTCTCACACCTTACCAAAACAGGCGACCTTTTTCAGCAGAGAGATGGATTTG 3000
Qy 3001 GGGCTACTCTCTCGCAGGCTCAGGTGGAAGATTCGTAGAGGAAACATTTAGTTTAAAGG 3060
Db 3001 GGGCTACTCTCTCGCAGGCTCAGGTGGAAGATTCGTAGAGGAAACATTTAGTTTAAAGG 3060

```
QY 3061 ACTTCATCCCTCCACCTATCCCTAACAGGCTGTAGATTACCAAAAAGAATTATTTTCAT 3120
Db 3061 ACTTCATCCCTCCACCTATCCCTAACAGGCTGTAGATTACCAAAAAGAATTATTTTCAT 3120
QY 3121 CACTAAAGAAATCTATTCACTCACTGCTCTTCAACAGACTTTTCTCTAGAGCCGCTCT 3180
Db 3121 CACTAAAGAAATCTATTCACTCACTGCTCTTCAACAGACTTTTCTCTAGAGCCGCTCT 3180
QY 3181 GCGTTTACTCTTGTGTTTCAAGGAGCACTTTTGTAAATCAATCATCTCTGTCACAGGCAG 3240
Db 3181 GCGTTTACTCTTGTGTTTCAAGGAGCACTTTTGTAAATCAATCATCTCTGTCACAGGCAG 3240
QY 3241 GAGGAGCTGATAATGAATCTTATTGAGCAGTATCTGATCCAAAGCGCTTCTCAGGCGG 3300
Db 3241 GAGGAGCTGATAATGAATCTTATTGAGCAGTATCTGATCCAAAGCGCTTCTCAGGCGG 3300
QY 3301 GCTTGAGTGAATCTGTACCTGAGTACAGTATCTTCTGTAATACATAAAACAAAAGC 3360
Db 3301 GCTTGAGTGAATCTGTACCTGAGTACAGTATCTTCTGTAATACATAAAACAAAAGC 3360
QY 3361 ATTTTGCTAAGGAGCACTAATATGATTTTAAAGTCTATGTTTAAATTAATATGTA 3420
Db 3361 ATTTTGCTAAGGAGCACTAATATGATTTTAAAGTCTATGTTTAAATTAATATGTA 3420
QY 3421 TTTTTCAGCTATTAGTGCATATATTTTATGGTGGGAATAAAATTTCTACTACAGAAAA 3480
Db 3421 TTTTTCAGCTATTAGTGCATATATTTTATGGTGGGAATAAAATTTCTACTACAGAAAA 3480
QY 3481 AAAAAAAAAAAAAAAAAAAAAA 3501
Db 3481 AAAAAAAAAAAAAAAAAAAAAA 3501
```

RESULT 6

```
ABS55045
ID ABS55045 standard; cDNA; 3501 BP.
AC ABS55045;
DT 09-DEC-2002 (first entry)
DE Human cDNA encoding receptor protein tyrosine kinase, FLK-2.
KW Human; ss; gene; FLK-2; foetal liver kinase; antianaemic;
KW ophthalmological; receptor protein tyrosine kinase; aplastic anaemia;
KW primitive haematopoietic cell; stem cell; macrocytic anaemia;
KW bone marrow damage; cancer chemotherapy.
OS Homo sapiens.
FH Key Location/Qualifiers
FT misc_feature 1..1689 a
FT CDS 58..3039
FT /note= "This region is specifically claimed in claim 72"
FT /tag= b
FT /product= "FLK-2"
FT /note= "This CDS (minus the stop codon) is specifically claimed in claim 73"
FT sig_peptide 58..138
FT /tag= c
FT mat_peptide 139..3036
FT /tag= d
FT /label= "Mature FLK_2"
FT /note= "This region is specifically claimed in claim 74"
US2002119545-A1.
29-AUG-2002.
01-JUN-2001; 2001US-0872136.
```

```
PR 19-NOV-1992; 92US-0977451.
PR 30-APR-1993; 93US-0055269.
PR 31-OCT-1994; 94US-0252498.
PR 15-FEB-1996; 96US-0601891.
PR 10-FEB-1998; 98US-0021324.
PR 10-DEC-1998; 98US-0208786.
PR 02-APR-1991; 91US-0679666.
PR 28-JUN-1991; 91US-0728913.
PR 15-NOV-1991; 91US-0793065.
PR 24-DEC-1991; 91US-0813593.
XX (LEMI/) LEMISCHKA I R.
XX Lemischka IR;
XX WPI; 2002-731356/79.
XX P-PSDB; ABG70916.
XX New protein tyrosine kinase expressed in primitive hematopoietic cells
XX (HC) and not expressed in mature HC, and ligands for the protein, stem
XX useful for stimulating proliferation of primitive hematopoietic stem
XX cells
XX Claim 9; Fig 1b; 64pp; English.
XX The invention relates to a receptor protein tyrosine kinase (pTK)
XX expressed in primitive haematopoietic cells and not expressed in mature
XX haematopoietic cells, and named FLK-1 and -2 (foetal liver kinase).
XX Also included are the nucleic acids encoding the FLK proteins, FLK
XX expression vectors, a ligand that binds to human or murine FLK-2 or FLK-1
XX and stimulates the proliferation and/or differentiation of the primitive
XX haematopoietic cells and a murine cell line 2018 having American Type
XX Culture Collection (ATCC) accession number ATCC CRL 10907. The ligands
XX are useful for stimulating the proliferation and/or differentiation of
XX primitive mammalian haematopoietic stem cells. The receptor pTK molecules
XX are useful for stimulating the self renewal of the totipotent
XX haematopoietic stem cell and to stimulate the development of all cells of
XX the haematopoietic system both in vitro and in vivo. The ligands for the
XX receptors act as haematopoietic growth factors. The ligands are useful in
XX treating humans whose primitive stem cells do not sufficiently undergo
XX self-renewal e.g. macrocytic and aplastic anaemia and bone marrow damage
XX resulting from cancer chemotherapy and radiation. The present
XX sequence encodes human FLK-2.
XX Sequence 3501 BP; 1068 A; 709 C; 784 G; 940 T; 0 other;
Query Match 100.0%; Score 3501; DB 24; Length 3501;
Best Local Similarity 100.0%; Pred. No 0;
Matches 3501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGAGGGCGGCATCCGAGGGCTGGGCGGCGGCCCTGGGGGACCCCGGGCTCCGAGGCCCATG 60
Db 1 CGAGGGCGGCATCCGAGGGCTGGGCGGCGGCCCTGGGGGACCCCGGGCTCCGAGGCCCATG 60
QY 61 CGGCGGTTGGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 120
Db 61 CGGCGGTTGGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 120
QY 121 TTTGGGACTATTACAATCAAGATCTGCTGTCATCAAGTGTGTTTAAATCAATCATTAAG 180
Db 121 TTTGGGACTATTACAATCAAGATCTGCTGTCATCAAGTGTGTTTAAATCAATCATTAAG 180
QY 181 AACATGATTTTCAGTGGGGAAGTCATCATCATATCCCATGTTATCAGAAATCCCGGAA 240
Db 181 AACATGATTTTCAGTGGGGAAGTCATCATCATATCCCATGTTATCAGAAATCCCGGAA 240
QY 241 GACCTCGGCTGTGCGTTGAGACCCAGAGCTCAGGAGCAGTGTACGAAGCTGCGGCTGTG 300
Db 241 GACCTCGGCTGTGCGTTGAGACCCAGAGCTCAGGAGCAGTGTACGAAGCTGCGGCTGTG 300
QY 301 GAAGTGAATGATCTGCTTCCATCACAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
Db 301 GAAGTGAATGATCTGCTTCCATCACAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
```


181 AACAAATGATTCATCAGTGGGAAGTCAATCATATATCCATGTTATCAGAAATCCCGGAA 240
181 AACAAATGATTCATCAGTGGGAAGTCAATCATATATCCATGTTATCAGAAATCCCGGAA 240
241 GACCTGGGTGTGGTGGAGACCCAGAGCTCAGGGAAGTGTAGAAAGTGGCGCTGTG 300
241 GACCTGGGTGTGGTGGAGACCCAGAGCTCAGGGAAGTGTAGAAAGTGGCGCTGTG 300
301 GAAGTGGATGATCTGCTTCCATCACACTGCAAGTGTGTCGATGCGCCAGGGAACATT 360
301 GAAGTGGATGATCTGCTTCCATCACACTGCAAGTGTGTCGATGCGCCAGGGAACATT 360
361 TCCTGTCTCTGGGTCTTTAAGCACAGCTCCCTGAAATGCCAGCCACATTTTGAATTA 420
361 TCCTGTCTCTGGGTCTTTAAGCACAGCTCCCTGAAATGCCAGCCACATTTTGAATTA 420
421 AACAGAGAGTGTGTTCCATGGTCATTTTGAATGAATGACAGAAACCCAGCTGGGAATAC 480
421 AACAGAGAGTGTGTTCCATGGTCATTTTGAATGAATGACAGAAACCCAGCTGGGAATAC 480
481 CTACTTTTATTCAGAGTGAAGCTACCAATTTACACAAATTTGTTTACAGTGAATAGA 540
481 CTACTTTTATTCAGAGTGAAGCTACCAATTTACACAAATTTGTTTACAGTGAATAGA 540
541 AATACCTGTCTTTACACATTAAGGAAGCTTTACTTTAGAAAAATGGAACCCAGAGGCC 600
541 AATACCTGTCTTTACACATTAAGGAAGCTTTACTTTAGAAAAATGGAACCCAGAGGCC 600
601 CTGCTGTGCATATCTGAGAGCGTTCCAGAGCGATCGTGGAAATGGGTCTTTGCGATTCA 660
601 CTGCTGTGCATATCTGAGAGCGTTCCAGAGCGATCGTGGAAATGGGTCTTTGCGATTCA 660
661 CAGGGGAAGCTGTAAGGAAGAAAGTCCAGCTGTTGTTTAAAGAGGAGAAAGTGCCTT 720
661 CAGGGGAAGCTGTAAGGAAGAAAGTCCAGCTGTTGTTTAAAGAGGAGAAAGTGCCTT 720
721 CATGAATTTATTTGGGACGACATAAGTGTGTCGACAGAAATGAATCTGGGACGGAATGC 780
721 CATGAATTTATTTGGGACGACATAAGTGTGTCGACAGAAATGAATCTGGGACGGAATGC 780
781 ACCAGGCTGTTTCAAAATAGATCTAAATCAAACTCCTCAGACCAATTCGACAAATTTT 840
781 ACCAGGCTGTTTCAAAATAGATCTAAATCAAACTCCTCAGACCAATTCGACAAATTTT 840
841 CTTAAAGTAGGGAAACCTTATGATAGTGAAGTGAAGCTGTTTCATGTGAACCATGGATTC 900
841 CTTAAAGTAGGGAAACCTTATGATAGTGAAGTGAAGCTGTTTCATGTGAACCATGGATTC 900
901 GGGCTCACCTGGGAATTTAGAAACAAAGCACTCAGAGGGCACTACTTTGAGATGAT 960
901 GGGCTCACCTGGGAATTTAGAAACAAAGCACTCAGAGGGCACTACTTTGAGATGAT 960
961 ACCTATTCAACAAACAGAACTATGATACGGATCTGTTGCTTTGTTATCATCAGTGGCA 1020
961 ACCTATTCAACAAACAGAACTATGATACGGATCTGTTGCTTTGTTATCATCAGTGGCA 1020
1021 AGAAACGACACCGGATACTACCTGTTTCCTTTCAAAGCATCCAGTCAATCAGCTTTG 1080
1021 AGAAACGACACCGGATACTACCTGTTTCCTTTCAAAGCATCCAGTCAATCAGCTTTG 1080
1081 GTTACCATCTAGGAAGGATTTATAAATGCTACCAATTCAGTGAAGATTTAGAAAT 1140
1081 GTTACCATCTAGGAAGGATTTATAAATGCTACCAATTCAGTGAAGATTTAGAAAT 1140
1141 GACCAATATGAAGTGTGTTTCTGTCAGGTTTAAAGCCTACCCACAAATCAGATGT 1200
1141 GACCAATATGAAGTGTGTTTCTGTCAGGTTTAAAGCCTACCCACAAATCAGATGT 1200
1201 ACCTGGAACCTTCTCTCGAAATCATTTCTTGTGAGCAAAAGGGTCTTGTAAACGGATAC 1260
1201 ACCTGGAACCTTCTCTCGAAATCATTTCTTGTGAGCAAAAGGGTCTTGTAAACGGATAC 1260
1261 AGCATATCCAAGTTTGGCAATCATAGCACACGAGGAGAGATATATATTCATGCGAGAA 1320

1261 AGCATATCCAAGTTTGGCAATCATAGCACACGAGGAGATATATATTCATGCGAGAA 1320
1321 AATGATGATGCCAATTTTACCAAAATGTTTCAAGTGAATATAAGAGGAACCTCAAGTG 1380
1321 AATGATGATGCCAATTTTACCAAAATGTTTCAAGTGAATATAAGAGGAACCTCAAGTG 1380
1381 CTGCGAGAAGCATCGGCAAGTCTGCGATGCGTCTGCGATGCGATGCGATGCGATGCGAT 1440
1381 CTGCGAGAAGCATCGGCAAGTCTGCGATGCGTCTGCGATGCGATGCGATGCGATGCGAT 1440
1441 TGGACCTGGGAAGAGTGTTCAGACAAAGTCTCCCACTCAGACAGAGAGATCAGAGAGGA 1500
1441 TGGACCTGGGAAGAGTGTTCAGACAAAGTCTCCCACTCAGACAGAGAGATCAGAGAGGA 1500
1501 GTCTGGGAATGAAGAGGTCTTAAAGAGTGTTCAGACAAAGTCTGAGCAGTACTCTA 1560
1501 GTCTGGGAATGAAGAGGTCTTAAAGAGTGTTCAGACAAAGTGTTCAGCAGTACTCTA 1560
1561 AACATGAGTGAAGCCTAATAAAGGTTCTGCTCAAGTCTGTCATCAATTTCCCTTGGC 1620
1561 AACATGAGTGAAGCCTAATAAAGGTTCTGCTCAAGTCTGTCATCAATTTCCCTTGGC 1620
1621 ACATCTTGTGAGAGCTCTTTTAACTCTCCAGCCCTTCCCTTTCATCCAGACACAC 1680
1621 ACATCTTGTGAGAGCTCTTTTAACTCTCCAGCCCTTCCCTTTCATCCAGACACAC 1680
1681 ATCTCATTTCTATGCAACAAATTTGTTGTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1740
1681 ATCTCATTTCTATGCAACAAATTTGTTGTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1740
1741 ATTTGTCAAGTACAAAGCAATTTAGTATGAAGCCAGCTACAGATGGTACAGGTG 1800
1741 ATTTGTCAAGTACAAAGCAATTTAGTATGAAGCCAGCTACAGATGGTACAGGTG 1800
1801 ACCGGCTCTCAGATAATGATGTTCTCTGCTGATTTTCAAGATATGAATATGATCTC 1860
1801 ACCGGCTCTCAGATAATGATGTTCTCTGCTGATTTTCAAGATATGAATATGATCTC 1860
1861 AAATGGGAGTTTCCAGAGAAAAATTTAGAGTTTGGGAAGGTTAGGATCAGGTCTCTTT 1920
1861 AAATGGGAGTTTCCAGAGAAAAATTTAGAGTTTGGGAAGGTTAGGATCAGGTCTCTTT 1920
1921 GGAAGATGATGAACGCAACAGCTTATGGAATTTAGCAAAAAGGAGTCTCAATCCAGTT 1980
1921 GGAAGATGATGAACGCAACAGCTTATGGAATTTAGCAAAAAGGAGTCTCAATCCAGTT 1980
1981 GCGTCAAAATGCTGGAAGAAAGCAGAGCTCTGAAAGAGAGGCTCATGTGAGAA 2040
1981 GCGTCAAAATGCTGGAAGAAAGCAGAGCTCTGAAAGAGAGGCTCATGTGAGAA 2040
2041 CTCAAGATGATGACCCAGCTGGGAAGCCAGAGAAATTTGGAACCTCTGGGGCGTGC 2100
2041 CTCAAGATGATGACCCAGCTGGGAAGCCAGAGAAATTTGGAACCTCTGGGGCGTGC 2100
2101 ACATGTCAGGACCAATTTTACTGATTTTGAATCTGTTGCTATGTTGATCTCTCAAC 2160
2101 ACATGTCAGGACCAATTTTACTGATTTTGAATCTGTTGCTATGTTGATCTCTCAAC 2160
2161 TATCTAAGAGTAAAGAGAAAAATTTTCAAGGCTTTGGAAGATTTTCAAGGAACAC 2220
2161 TATCTAAGAGTAAAGAGAAAAATTTTCAAGGCTTTGGAAGATTTTCAAGGAACAC 2220
2221 AATTTCAAGTTTACCCCACTTTCCCAATCATCTCAAAATTCAGCATGCTGGTTCAAGA 2280
2221 AATTTCAAGTTTACCCCACTTTCCCAATCATCTCAAAATTCAGCATGCTGGTTCAAGA 2280
2281 GAAAGTTTCAAGTACACCCGAGCTCGGATCAATCTCAGGGCTTTCATGGGAATTCATTTCAC 2340
2281 GAAAGTTTCAAGTACACCCGAGCTCGGATCAATCTCAGGGCTTTCATGGGAATTCATTTCAC 2340
2341 TCTGAAGATGAATTTGAATGAAAAACCAAAAGGCTGGAAGAGGAGGACTTGAAT 2400

Db 2341 TCTGAAGATGAATTGAATATGATAAACAACAAAAGGCTGGAGAGGAGCATTGAT 2400
Qy 2401 GTGCTTACATTGAAGATCTCTTTGCTTTCATATCAAGTTGCCAAAGGAATGAAATTT 2460
Db 2401 GTGCTTACATTGAAGATCTCTTTGCTTTCATATCAAGTTGCCAAAGGAATGAAATTT 2460
Qy 2461 CTGGAATTTAAGTCGTGTGTTTCAGAGAGACCTGGCCGCGCAGGAACGTTGTTCACCCAC 2520
Db 2461 CTGGAATTTAAGTCGTGTGTTTCAGAGAGACCTGGCCGCGCAGGAACGTTGTTCACCCAC 2520
Qy 2521 GGGAAAGTGGTGAAGATATGTGATCTTTGGATTGGCTCGAGATATCATGAGTGATTTCAAC 2580
Db 2521 GGGAAAGTGGTGAAGATATGTGATCTTTGGATTGGCTCGAGATATCATGAGTGATTTCAAC 2580
Qy 2581 TATGTTGTCAGGGGCAATGCCCGTCTGCTCTGTAATAATGGATGGCCCGCAAGCCTGTTT 2640
Db 2581 TATGTTGTCAGGGGCAATGCCCGTCTGCTCTGTAATAATGGATGGCCCGCAAGCCTGTTT 2640
Qy 2641 GAAGGCATCTACACCATTAAGAGTGATGTCGTGTCATATGGAATATTACTGTGGGAATC 2700
Db 2641 GAAGGCATCTACACCATTAAGAGTGATGTCGTGTCATATGGAATATTACTGTGGGAATC 2700
Qy 2701 TTCTCACTTGGTGGAATCCTTACCTGGCATTCGGTTGATGCTAACTTCTACAAACTG 2760
Db 2701 TTCTCACTTGGTGGAATCCTTACCTGGCATTCGGTTGATGCTAACTTCTACAAACTG 2760
Qy 2761 ATTCAAATGGAATTTAAATGGATCAGCCATTTATGCTACAGAGAAATATACATTATA 2820
Db 2761 ATTCAAATGGAATTTAAATGGATCAGCCATTTATGCTACAGAGAAATATACATTATA 2820
Qy 2821 ATGCAATCTGCTGGCTTTTGACTCAAGGAAGAGGCCATCTTCCCTAAATTTGACTTCG 2880
Db 2821 ATGCAATCTGCTGGCTTTTGACTCAAGGAAGAGGCCATCTTCCCTAAATTTGACTTCG 2880
Qy 2881 TTTTATGATGTCAGTGGCAGATGCAAGAAGCGATGATATCAGATGTCGATGGCCGT 2940
Db 2881 TTTTATGATGTCAGTGGCAGATGCAAGAAGCGATGATATCAGATGTCGATGGCCGT 2940
Qy 2941 GTTTCGGAATGTCCTCACACCTTACCAAAAGGCGACCTTTTTCAGCAGAGATGGAATTG 3000
Db 2941 GTTTCGGAATGTCCTCACACCTTACCAAAAGGCGACCTTTTTCAGCAGAGATGGAATTG 3000
Qy 3001 GGGCTACTCTCTCGCAGGCTCAGTCAAGATTCGTAGAGAAACAATTTAGTTTAAAG 3060
Db 3001 GGGCTACTCTCTCGCAGGCTCAGTCAAGATTCGTAGAGAAACAATTTAGTTTAAAG 3060
Qy 3061 ACTTCATCCCTCCACCTATCCCTAACAGGCTGTAGATTACCAAAACAGATTAATTTTCAT 3120
Db 3061 ACTTCATCCCTCCACCTATCCCTAACAGGCTGTAGATTACCAAAACAGATTAATTTTCAT 3120
Qy 3121 CACTAAAAGAAAATCTATTATCACTGCTGCTTCCACAGACTTTTCTCTAGAGCCGCT 3180
Db 3121 CACTAAAAGAAAATCTATTATCACTGCTGCTTCCACAGACTTTTCTCTAGAGCCGCT 3180
Qy 3181 CGGTTTACTCTGTTTCAAGGGACTTTTGTAAATCAATCATCTCTGTGCACAGGCG 3240
Db 3181 CGGTTTACTCTGTTTCAAGGGACTTTTGTAAATCAATCATCTCTGTGCACAGGCG 3240
Qy 3241 GAGGAGCTGATAATCAACTTTATTGGACATTCATCTGCATCCAGGCGCTCTCAGGCG 3300
Db 3241 GAGGAGCTGATAATCAACTTTATTGGACATTCATCTGCATCCAGGCGCTCTCAGGCG 3300
Qy 3301 GCTTGAGTGAATTTGTTACCTGAAGTACAGATATTTCTTTGTAATACATAAAACAAAGC 3360
Db 3301 GCTTGAGTGAATTTGTTACCTGAAGTACAGATATTTCTTTGTAATACATAAAACAAAGC 3360
Qy 3361 ATTTTGTACAGAGAACTAATATGATTTTAAAGTCTATGTTTAAATATATATGTA 3420
Db 3361 ATTTTGTACAGAGAACTAATATGATTTTAAAGTCTATGTTTAAATATATATGTA 3420
Qy 3421 TTTTTCAGCTATTAGTATATATTTTATGGTGGGAATTAATTTCTACTACAGAAAA 3480
Db 3421 TTTTTCAGCTATTAGTATATATTTTATGGTGGGAATTAATTTCTACTACAGAAAA 3480

Qy 3481 AAAAAAAAAAAAAAAAAAAAAA 3501
Db 3481 AAAAAAAAAAAAAAAAAAAAAA 3501

RESULT 8

AAQ40915

ID AAQ40915 standard; cDNA; 3501 BP.

XX AAQ40915;

XX 25-MAR-2003 (updated)

DT 19-OCT-1993 (first entry)

DE Human flk-2 cDNA.

XX Murine; receptor; protein; tyrosine kinase; ptk; flk-2; primitive;
XX hematopoietic cell; mature; family; conserved; region;
XX catalytic domain; c-kit; fetal liver kinase; flk; fetal; spleen;
XX thymus; adult; brain; bone marrow; multipotential; CFU-Blast colony;
XX hierarchy; transduction; T-lymphoid; lineage; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

FH CDS 58..3039

FT /*tag= a

FT sig_peptide 58..138

FT /*tag= b

FT /notes= "Hydrophobic leader sequence"

FT mat_peptide 139..3036

FT /*tag= c

XX WO9310136-A1.

XX 27-MAY-1993.

XX 16-NOV-1992; 92WO-US09893.

XX 15-NOV-1991; 91US-0793065.

XX (UYPR-) UNIV PRINCETON.

XX Lemischka IR;

XX WPI: 1993-182479/22.

XX P-PSDB; AAR37503.

XX Totipotent haematopoietic stem cell receptors, their ligands and
XX DNA sequences - for treating anaemia(s) and bone marrow damage
XX due to e.g. cancer chemotherapy or radiotherapy

XX Claim 9; Fig 1b; 127pp; English.

XX This sequence encodes the human receptor protein tyrosine kinase
XX (ptk), flk-2. This nucleic acid is expressed in primitive hemato-
XX poietic cells and not in mature hematopoietic cells. Members of
XX this family of ptk's can be recognised by the conserved amino acid
XX regions in the catalytic domain. This family of ptk's also contains
XX c-kit. These new receptors are termed fetal liver kinases (flk's)
XX after the tissue in which they were discovered. flk-2 is also
XX expressed in fetal spleen, fetal thymus, adult brain and adult
XX bone marrow. flk-2 is expressed in individual multipotential CFU-
XX Blast colonies capable of generating numerous multilineage colonies
XX upon replating. It is likely therefore, that flk-2 is expressed in
XX the entire primitive portion of the hematopoietic hierarchy. This is
XX consistent with flk-2 being important in transducing putative self-
XX renewal signals from the environment. flk-2 is the first receptor
XX ptk known to be expressed in the T-lymphoid lineage.
XX (Updated on 25-MAR-2003 to correct PN field.)

XX Sequence 3501 BP; 1068 A; 708 C; 785 G; 940 T; 0 other;

Query Match	100.0%; Score 3499.4; DB 14; Length 3501;
Best Local Similarity	100.0%; Pred. No. 0;
Matches 3500; Conservative	0; Mismatches 1; Indels 0; Gaps 0;
Qy	1 CGAGCGGCATCCGAGGGCTGGCGCGCCCTGGGGACCCCGGGCTCCGGAGCCCATG 60
Db	
Qy	1 CGAGCGGCATCCGAGGGCTGGCGCGCCCTGGGGACCCCGGGCTCCGGAGCCCATG 60
Db	
Qy	61 CCGGCGTTGGCGCGACGCGGGGACCCGCTGCTGCTGTTGTTTTTCTGCAATGATA 120
Db	
Qy	121 TTTGGGACTATTACAAATCAAGATCTGCTGTGATCAAGTGTGTTTTTAATCAATATAAG 180
Db	
Qy	181 AACATGATTCATCAGTGGGAGTTCATCATCATATCCATGGTATCAGATCCCCGAA 240
Db	
Qy	241 GACCTCGGCTGTGCTTGAGACCCAGAGCTCAGGCACAGTGTACGAAGCTGCCGCTGTG 300
Db	
Qy	301 GAAGTGAATGATCTGCTTCATCACACTGCAAGTGTGTCGATGCCCGAGGAAACATT 360
Db	
Qy	361 TCCTGTCTCGGCTTTTAAGCAGAGCTCCCTGTAATTGCCAGCACATTTTGTATTCAA 420
Db	
Qy	421 AACAGAGAGTGTGTTCCATGGTCAATTTTGAATAATGACAGAAACCCAGCTGGAGATAC 480
Db	
Qy	481 CTACTTTTTATTTCAGAGTGAAGCTACCAATTACAAATATTGTTTACAGTGAGTATAAGA 540
Db	
Qy	541 AATACCTGCTTTACATTAAGAAGACCTTTACTTTAGAAAAATGGAACACCGAGCGCC 600
Db	
Qy	601 CTGGCTGCAATCTGAGAGCGTTCCAGAGCCGATCGTGAATGGGTGCTTTGGCATCA 660
Db	
Qy	661 CAGGGGAAAGCTGTAAAGAAAGATCCAGCTGTTGTTAAAAAGGAGGAAAAAGTGCCT 720
Db	
Qy	721 CATGAATTAATTTGGGACGGACATAAGGTGCTGTGCCAGAAATGAATGCGGACGGGAATGC 780
Db	
Qy	781 ACCAGGCTGTTCAAAATAGATCTTAATCAAACTCCTCAGACCAATTTGCCACAAATTAAT 840
Db	
Qy	841 CTTAAAGTAGGGGAAACCTTATGATTAAGGTGCAAGCTGTTCAATGTGAACCAATGATTC 900
Db	
Qy	901 GGGCTCACCTGGGAATTAGAAAACAAAGCACTCAGGAGGGCAACTACTTTTGATGATGAT 960
Db	
Qy	961 ACCTATTCAACAAACAGAACTATGATACGGATTCGTTTGTCTTTGTGTATCATACAGTGGCA 1020
Db	

CC to be uncommitted. Therefore, thymocytes expressing flk-2 may be
CC multipotential. flk-2 is the first receptor tyrosine kinase known to
CC be expressed in the T-lymphoid lineage.
CC (Updated on 25-MAR-2003 to correct PN field.)
XX

SQ Sequence 3501 BP; 1067 A; 708 C; 786 G; 940 T; 0 other;

Query Match 99.9%; Score 3497.8; DB 14; Length 3501;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3499; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY	1	CGAGCGGCATCCGAGGCTGGCGCGCCCTGGGGACCCCGGCTCGGAGGCCATG	60
DB	1	CGAGCGGCATCCGAGGCTGGCGCGCCCTGGGGACCCCGGCTCGGAGGCCATG	60
QY	61	CCGGCGTTGGCGCGACCGCGCGCCCTGGCGCTGCTGTTGTTTTCTGCAATGATA	120
DB	61	CCGGCGTTGGCGCGACCGCGCGCCCTGGCGCTGCTGTTGTTTTCTGCAATGATA	120
QY	121	TTTGGGACTATTACAAATCAAGATCTGCTGTGATCAAGTGTGTTTTAAATCAATTAAG	180
DB	121	TTTGGGACTATTACAAATCAAGATCTGCTGTGATCAAGTGTGTTTTAAATCAATTAAG	180
QY	181	AACATGATTCATCAGTGGGAGTCATCATATCCCATGTCAGATCCCGGAA	240
DB	181	AACATGATTCATCAGTGGGAGTCATCATATCCCATGTCAGATCCCGGAA	240
QY	241	GACCTCGGCTGTGAGACCCAGAGCTCAGGACAGTGTACGAGCTGCCCTGTG	300
DB	241	GACCTCGGCTGTGAGACCCAGAGCTCAGGACAGTGTACGAGCTGCCCTGTG	300
QY	301	GAAGTGGATGATCTGCTTCCATCAGTGCAGTGGTGGTCCCGGGAACATT	360
DB	301	GAAGTGGATGATCTGCTTCCATCAGTGCAGTGGTGGTCCCGGGAACATT	360
QY	361	TCCTGCTCTGGCTTTAAGCAGAGCTCCCTGGAATGCGCCACATTTGATTACA	420
DB	361	TCCTGCTCTGGCTTTAAGCAGAGCTCCCTGGAATGCGCCACATTTGATTACA	420
QY	421	AACAGAGAGTGTGTTTCCATGTCATTTGAAATGACAGAACCCAGCTGGAGATAC	480
DB	421	AACAGAGAGTGTGTTTCCATGTCATTTGAAATGACAGAACCCAGCTGGAGATAC	480
QY	481	CTACTTTTTTATTCAGATGGAAGTACCAATACAAATATTTTACAGTGAATAGA	540
DB	481	CTACTTTTTTATTCAGATGGAAGTACCAATACAAATATTTTACAGTGAATAGA	540
QY	541	AATACCTGCTTTACATTAAGAGACCTTACTTTAGAAAAATGAAAAACAGGCGCC	600
DB	541	AATACCTGCTTTACATTAAGAGACCTTACTTTAGAAAAATGAAAAACAGGCGCC	600
QY	601	CTGCTGTCATATCTGAGAGCTTCCAGAGCGATCGTGGATGGTGCTTGGATTCA	660
DB	601	CTGCTGTCATATCTGAGAGCGTTCAGAGCGATCGTGGATGGTGCTTGGATTCA	660
QY	661	CAGGGGAAAGCTGTAAAGAGAAAGTCCAGCTGTTGTTAAAAAGGAGAAAAAGTGCTT	720
DB	661	CAGGGGAAAGCTGTAAAGAGAAAGTCCAGCTGTTGTTAAAAAGGAGAAAAAGTGCTT	720
QY	721	CATGAATATTTGGGACCGGACATAGGTCGTGTCAGAAATGAATCGGGCAGGGAATGC	780
DB	721	CATGAATATTTGGGACCGGACATAGGTCGTGTCAGAAATGAATCGGGCAGGGAATGC	780
QY	781	ACAGGCTGTTCAATAGATCTAAATCAAACTCTCAGACCATTTGCCAATATTTT	840
DB	781	ACAGGCTGTTCAATAGATCTAAATCAAACTCTCAGACCATTTGCCAATATTTT	840
QY	841	CTTAAAGTGGGAAACCTTATGATAGGTGCAAGCTGTTCAATGCAATGATTC	900
DB	841	CTTAAAGTGGGAAACCTTATGATAGGTGCAAGCTGTTCAATGCAATGATTC	900
QY	901	GGGCTCACCTGGGAATTAGAAAAAAGCACTCGAGGAGGCAACTACTTTGAGATGAT	960
DB	901	GGGCTCACCTGGGAATTAGAAAAAAGCACTCGAGGAGGCAACTACTTTGAGATGAT	960

DB	901	GGGCTCACCTGGGAATTAGAAAAAAGCACTCGAGGAGGCAACTACTTTGAGATGAT	960
QY	961	ACCTATTCAACAAAACAGACTATGATACGGATCTGTTTCTTTGATCATCAGTGGCA	1020
DB	961	ACCTATTCAACAAAACAGACTATGATACGGATCTGTTTCTTTGATCATCAGTGGCA	1020
QY	1021	AGAAACGACACCGGATACACACTTGTCTCTTCAAGCATCCAGTCAATCAGCTTTG	1080
DB	1021	AGAAACGACACCGGATACACACTTGTCTCTTCAAGCATCCAGTCAATCAGCTTTG	1080
QY	1081	GTTACATCTGTAAGAAAGGATTTAAATGCTACCAATTTCAAGTGAAGATTTGAAAT	1140
DB	1081	GTTACATCTGTAAGAAAGGATTTAAATGCTACCAATTTCAAGTGAAGATTTGAAAT	1140
QY	1141	GACCAATATGAAGAGTGTGTTTTTCTGTCAAGTTTAAAGCCCTACCAAAATCAGAT	1200
DB	1141	GACCAATATGAAGAGTGTGTTTTTCTGTCAAGTTTAAAGCCCTACCAAAATCAGAT	1200
QY	1201	ACGTGGACCTTCTCTCGAAATCATTTCTGTGAGCAAAAGGCTCTTCATACCGATAC	1260
DB	1201	ACGTGGACCTTCTCTCGAAATCATTTCTGTGAGCAAAAGGCTCTTCATACCGATAC	1260
QY	1261	AGCATATCAAAGTTTGAATCAATCAAGCAGCAGGAGAAATATATATTCATGCGAA	1320
DB	1261	AGCATATCAAAGTTTGAATCAATCAAGCAGCAGGAGAAATATATATTCATGCGAA	1320
QY	1321	AATGATGATGCCAATTTACCAAAATGTTTCAAGTGAATATAAGAAAGAAACCTCAAGTG	1380
DB	1321	AATGATGATGCCAATTTACCAAAATGTTTCAAGTGAATATAAGAAAGAAACCTCAAGTG	1380
QY	1381	CTCGCAGAGCATCGGCAAGTCTCAGCAAGTCTCCAACTGCACAGAGAGATCACAGAGA	1440
DB	1381	CTCGCAGAGCATCGGCAAGTCTCAGCAAGTCTCCAACTGCACAGAGAGATCACAGAGA	1440
QY	1441	TGGACCTGGAAGAGTGTTCAGCAAGTCTCCAACTGCACAGAGAGATCACAGAGA	1500
DB	1441	TGGACCTGGAAGAGTGTTCAGCAAGTCTCCAACTGCACAGAGAGATCACAGAGA	1500
QY	1501	GTCTGGAATAGAAAGGCTTAAAGAAAGTGTGAGCAGTGGTCTCGAGCAGTACTCTA	1560
DB	1501	GTCTGGAATAGAAAGGCTTAAAGAAAGTGTGAGCAGTGGTCTCGAGCAGTACTCTA	1560
QY	1561	AACATGATGAGCAGCATAAAGGCTTCTGTGCAAGTCTGTGCAATATTCCTTGGC	1620
DB	1561	AACATGATGAGCAGCATAAAGGCTTCTGTGCAAGTCTGTGCAATATTCCTTGGC	1620
QY	1621	ACATCTTGTGAGCAGTCTTTTAAACCTCCAGGCCCTTCCCTTTCATCCAGACAAC	1680
DB	1621	ACATCTTGTGAGCAGTCTTTTAAACCTCCAGGCCCTTCCCTTTCATCCAGACAAC	1680
QY	1681	ATCTCATTTAAGCAAAATTTGTTGTTCTCTCTCTTCAATGTCGTTTAAACCTGCTA	1740
DB	1681	ATCTCATTTAAGCAAAATTTGTTGTTCTCTCTCTTCAATGTCGTTTAAACCTGCTA	1740
QY	1741	ATTTGTCAAGTACAAAGCAATTTAGGTATGAAGCCAGCTACAGATGGTACAGTG	1800
DB	1741	ATTTGTCAAGTACAAAGCAATTTAGGTATGAAGCCAGCTACAGATGGTACAGTG	1800
QY	1801	ACCGCTCTCAGATAATGATCTTCTACGTGATTTTCAAGAAATATGATATGATCTC	1860
DB	1801	ACCGCTCTCAGATAATGATCTTCTACGTGATTTTCAAGAAATATGATATGATCTC	1860
QY	1861	AAATGGAGTTTCCAAGAGAAAAATTTAGAGTTTGGGAAAGTACTAGGATCAGGTCTTTT	1920
DB	1861	AAATGGAGTTTCCAAGAGAAAAATTTAGAGTTTGGGAAAGTACTAGGATCAGGTCTTTT	1920
QY	1921	GGAAAGTATGACGCAACAGCTTATGGAATTTAGCAAAAACAGGAGTCTCAATCCAGTT	1980
DB	1921	GGAAAGTATGACGCAACAGCTTATGGAATTTAGCAAAAACAGGAGTCTCAATCCAGTT	1980
QY	1981	GCCTGCAAAATGCTGAAAGAAAGCAGCAGCTCTGAAAGAGAGGCACTCATGTCAGAA	2040
DB	1981	GCCTGCAAAATGCTGAAAGAAAGCAGCAGCTCTGAAAGAGAGGCACTCATGTCAGAA	2040

Qy	2041	CTCAAGATGATGACCCAGCTGGGAAGCCACGAGAAATATTGTGAACTCTGCTGGGGCGTG	2100
Db	2041	CTCAAGATGATGACCCAGCTGGGAAGCCACGAGAAATATTGTGAACTCTGCTGGGGCGTG	2100
Qy	2101	ACACTGTCAAGACCAATTTACTTTGAATCTGTTTGAATCTGTTGCTATGGTGATCTTCTCAAC	2160
Db	2101	ACACTGTCAAGACCAATTTACTTTGAATCTGTTTGAATCTGTTGCTATGGTGATCTTCTCAAC	2160
Qy	2161	TATCTAAGAAGTAAAAAGAGAAAAATTTCAAGGACTTTGGAACAGAGATTTTCAAGGAACAC	2220
Db	2161	TATCTAAGAAGTAAAAAGAGAAAAATTTCAAGGACTTTGGAACAGAGATTTTCAAGGAACAC	2220
Qy	2221	AAATTCAGTTTTTACCCCACTTTCCCAATCACATCCAAATTCAGAGCATGCTGTTCAAGA	2280
Db	2221	AAATTCAGTTTTTACCCCACTTTCCCAATCACATCCAAATTCAGAGCATGCTGTTCAAGA	2280
Qy	2281	GAAGTTCAAGATACACCGGACTCGGATCAATCTCAGGGCTTCATGGGAATTCATTTCAAC	2340
Db	2281	GAAGTTCAAGATACACCGGACTCGGATCAATCTCAGGGCTTCATGGGAATTCATTTCAAC	2340
Qy	2341	TCCTGAAGATGAAATTTGAATATGAAAAACCAAAAAAGGCTGGAAGACAGAGACTTGAAT	2400
Db	2341	TCCTGAAGATGAAATTTGAATATGAAAAACCAAAAAAGGCTGGAAGACAGAGACTTGAAT	2400
Qy	2401	GTGCTTACATTTGAAGATCTTCTTTGCTTTGCATATCAAGTTGCCAAAAGATGGAATTT	2460
Db	2401	GTGCTTACATTTGAAGATCTTCTTTGCTTTGCATATCAAGTTGCCAAAAGATGGAATTT	2460
Qy	2461	CTGGAAATTTAAGTCGTGTGTTTCACAGAGACCTGGCCGCCAGGAACTGCTGTGCACCCAC	2520
Db	2461	CTGGAAATTTAAGTCGTGTGTTTCACAGAGACCTGGCCGCCAGGAACTGCTGTGCACCCAC	2520
Qy	2521	GGGAAAGTGTTGAAGATATGTGACTTTTGGATTTGGCTCGAGATATCATGTAGTATCCCAAC	2580
Db	2521	GGGAAAGTGTTGAAGATATGTGACTTTTGGATTTGGCTCGAGATATCATGTAGTATCCCAAC	2580
Qy	2581	TATGTTGTCAAGGGCAATGCCGTCGCTGTAATAATGGAATGSCCCCGAAGACCTGTTT	2640
Db	2581	TATGTTGTCAAGGGCAATGCCGTCGCTGTAATAATGGAATGSCCCCGAAGACCTGTTT	2640
Qy	2641	GAAGGCATCTACACCATTAAGATGATGTCTGGTTCATATGGAATATTACTGTGGGAAATC	2700
Db	2641	GAAGGCATCTACACCATTAAGATGATGTCTGGTTCATATGGAATATTACTGTGGGAAATC	2700
Qy	2701	TTCTCATTGGTGTGAATCCCTTACCCCTGGCAATCCGGTTGAGTGTAACCTCTCAAAACTG	2760
Db	2701	TTCTCATTGGTGTGAATCCCTTACCCCTGGCAATCCGGTTGAGTGTAACCTCTCAAAACTG	2760
Qy	2761	ATTCAAAAATGATTTAAAAATGGATCAGCCATTTTATGCTACAGAGAAATATACATTATA	2820
Db	2761	ATTCAAAAATGATTTAAAAATGGATCAGCCATTTTATGCTACAGAGAAATATACATTATA	2820
Qy	2821	ATGCAATCCTCTGGGCTTTTGACTCAAGGAAACGGCCATCCTTCCCTAAATTTGACTTCG	2880
Db	2821	ATGCAATCCTCTGGGCTTTTGACTCAAGGAAACGGCCATCCTTCCCTAAATTTGACTTCG	2880
Qy	2881	TTTTTTPAGGATGTACGCTGGCAGATGCAAGAACGATGTATCAGAATGTGGATGGCCGT	2940
Db	2881	TTTTTTPAGGATGTACGCTGGCAGATGCAAGAACGATGTATCAGAATGTGGATGGCCGT	2940
Qy	2941	GTTTCGGAAATGCTCACAACCTACCAAAACAGGCGACTTTTCAGCAGAGATGGAATTCG	3000
Db	2941	GTTTCGGAAATGCTCACAACCTACCAAAACAGGCGACTTTTCAGCAGAGATGGAATTCG	3000
Qy	3001	GGGCTACTCTCTCCGAGGCTCAGGTCGAAGATTCGTAGAGAAACAATTTAGTTTTAAGG	3060
Db	3001	GGGCTACTCTCTCCGAGGCTCAGGTCGAAGATTCGTAGAGAAACAATTTAGTTTTAAGG	3060
Qy	3061	ACTTCATCCCTCCACTATCCCTAACAGGCTGTAGATTTACCAAAACAAGATTAATTTTCAT	3120
Db	3061	ACTTCATCCCTCCACTATCCCTAACAGGCTGTAGATTTACCAAAACAAGATTAATTTTCAT	3120

Qy	3121	CAC	TAAAGAA	AATCTATTATCA	CTGCTGCTTCACCAGACTTTTCTCTAGAACCGTCT	3180
Db	3121	CAC	TAAAGAA	AATCTATTATCA	CTGCTGCTTCACCAGACTTTTCTCTAGAACCGTCT	3180
Qy	3181	GC	CTTTACTCT	TGTTTTCA	AGGAGCTTTTGTAAATCAAAATCATCTGTCACAAGGCAG	3240
Db	3181	GC	CTTTACTCT	TGTTTTCA	AGGAGCTTTTGTAAATCAAAATCATCTGTCACAAGGCAG	3240
Qy	3241	GAG	GAGCTGATA	TAACTTTATT	TGGAGCATTGATCTGCATCCAAGGCCCTTCTCAGGCCG	3300
Db	3241	GAG	GAGCTGATA	TAACTTTATT	TGGAGCATTGATCTGCATCCAAGGCCCTTCTCAGGCCG	3300
Qy	3301	GCT	TGAGTGA	ATTTGTG	TACTGAAGTACAGTATATCTTTGTAAATACATAAAACAAAGC	3360
Db	3301	GCT	TGAGTGA	ATTTGTG	TACTGAAGTACAGTATATCTTTGTAAATACATAAAACAAAGC	3360
Qy	3361	AT	TTTCTGCTA	GAGGAGAGCTA	ATATGATTTTTTAAAGTCTATGTTTTTAAATAAATATGTAAA	3420
Db	3361	AT	TTTCTGCTA	GAGGAGAGCTA	ATATGATTTTTTAAAGTCTATGTTTTTAAATAAATATGTAAA	3420
Qy	3421	TTTT	TCAGCTATT	TTTAGT	GATATATTTTATGGGTGGGAATAAAATTTCTACTACAGAAAA	3480
Db	3421	TTTT	TCAGCTATT	TTTAGT	GATATATTTTATGGGTGGGAATAAAATTTCTACTACAGAAAA	3480
Qy	3481	AAA	AAAAA	AAAAA	AAAAA 3501	
Db	3481	AAA	AAAAA	AAAAA	AAAAA 3501	
RESULT 10						
AAQ91536						
ID	AAQ91536	standard; DNA; 3476 BP.				
XX	AAQ91536;					
AC	AAQ91536;					
XX	25-MAR-2003 (updated)					
DT	29-DEC-1995 (first entry)					
XX	Human STK-1 cDNA.					
XX	STK-1; receptor PTK; protein tyrosine kinase; ss.					
OS	Homo sapiens.					
XX	Key	Location/Qualifiers				
FT	CDS	58..2976				
FT		/*tag= a				
FT	misc_feature	1..57				
FT		/*tag= b				
FT		/label= claimed antisense STK-1 oligo				
FT	misc_feature	121..148				
FT		/*tag= C				
FT		/label= see above				
XX	WO9519175-A1.					
PN	20-JUL-1995.					
PD	06-JAN-1995;	95WO-US00176.				
PF	14-JAN-1994;	94US-0183211.				
XX	(UWJO) UNIV JOHNS HOPKINS.					
PA	(UYPE-) UNIV PENNSYLVANIA.					
XX	Civin CI, Gewirtz AM, Small D;					
PI	WPI: 1995-263709/34.					
DR	P-PSDB; AAR75961.					
DR	Artificial STK-1 gene and gene-specific anti-sense oligo-nucleotide -					
XX	used to treat neoplastic diseases and as bone marrow purging agents for					
PT	treating leukaemia and neoplasia					

Claim 8; Fig 1; 66pp; English.
 The STK-1 gene encodes a receptor PTK which is expressed in proliferating hematopoietic stem cells but not in quiescent stem cells. The STK-1 gene is also expressed in certain malignant cells of non-hematopoietic origin. An antisense oligo specific for STK-1 is an oligo having a sequence (i) capable of forming a stable triplex with a portion of the STK-1 gene, or (ii) capable of forming a stable duplex with a portion of an mRNA transcript of the STK-1 gene. Antisense oligos capable of forming a stable duplex with a portion of a STK-1 mRNA transcript are given in AAQ91536 PT and in AAQ91537 and AAQ91538. The antisense oligos of the invention are useful in the treatment of hematologic malignancies characterized by STK-1 expression.
 CC (Updated on 25-MAR-2003 to correct PN field.)
 XX
 SQ Sequence 3476 BP; 1044 A; 709 C; 783 G; 940 T; 0 other;
 Query Match 99.2%; Score 3474.4; DB 16; Length 3476;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 3475; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 CGAGGCGGCATCCGAGGCTGGGCGGCGCCCTCGGGGACCCCGGCTCCGAGGCCATG 60
 DB |||||
 QY 1 CGAGGCGGCATCCGAGGCTGGGCGGCGCCCTCGGGGACCCCGGCTCCGAGGCCATG 60
 DB |||||
 QY 61 CCGCGGTGGCGCGACGCGGGGACCGTGCCTGCTCGTGTGTTTCTGCAATGATA 120
 DB |||||
 QY 61 CCGCGGTGGCGCGACGCGGGGACCGTGCCTGCTCGTGTGTTTCTGCAATGATA 120
 DB |||||
 QY 121 TTTGGGACTATTACAAATCAAGATCTGCCTGTGATCAAGTGTGTTTAAATCAATCATAAG 180
 DB |||||
 QY 121 TTTGGGACTATTACAAATCAAGATCTGCCTGTGATCAAGTGTGTTTAAATCAATCATAAG 180
 DB |||||
 QY 181 AACAAATGATTATCAGTGGGGAAGTCAATCATATATCCCATGGPATCAGAAATCCCGGAA 240
 DB |||||
 QY 181 AACAAATGATTATCAGTGGGGAAGTCAATCATATATCCCATGGPATCAGAAATCCCGGAA 240
 DB |||||
 QY 241 GACCTCGGTGTGCTTTGACAGCCAGAGCTCAGGACAGTGTACGAAGCTGCGGCTGTG 300
 DB |||||
 QY 241 GACCTCGGTGTGCTTTGACAGCCAGAGCTCAGGACAGTGTACGAAGCTGCGGCTGTG 300
 DB |||||
 QY 301 GAAGTGGATGATATCTGCTTCATCACAAGTGTGCTGATGATGCTCCAGGGAAACATT 360
 DB |||||
 QY 301 GAAGTGGATGATATCTGCTTCATCACAAGTGTGCTGATGATGCTCCAGGGAAACATT 360
 DB |||||
 QY 361 TCTGTCTCTGGGTCTTTAAGACAGCTCCCTGAAATGCGAGCCACATTTTGATTACAA 420
 DB |||||
 QY 361 TCTGTCTCTGGGTCTTTAAGACAGCTCCCTGAAATGCGAGCCACATTTTGATTACAA 420
 DB |||||
 QY 421 AACAGAGGAGTGTCTTCCATGTCATTTTGAAATGACAGAAACCCAGCTGAGGAATAC 480
 DB |||||
 QY 421 AACAGAGGAGTGTCTTCCATGTCATTTTGAAATGACAGAAACCCAGCTGAGGAATAC 480
 DB |||||
 QY 481 CTACTTTTATTTCAGAGTGAAGCTACCAATTAACAATATTGTTTACAGTGAATTAAGA 540
 DB |||||
 QY 481 CTACTTTTATTTCAGAGTGAAGCTACCAATTAACAATATTGTTTACAGTGAATTAAGA 540
 DB |||||
 QY 541 AATACCTGCTTTACACATTTAAGAAGACCTTACTTTAGAAAAATGAAAAACAGGACGCC 600
 DB |||||
 QY 541 AATACCTGCTTTACACATTTAAGAAGACCTTACTTTAGAAAAATGAAAAACAGGACGCC 600
 DB |||||
 QY 601 CTGGTCTGATATCTCAGAGCGTTCCAGAGCCGATCGTGGAAATGGGTGCTTTGCGATTCA 660
 DB |||||
 QY 601 CTGGTCTGATATCTCAGAGCGTTCCAGAGCCGATCGTGGAAATGGGTGCTTTGCGATTCA 660
 DB |||||
 QY 661 CAGGGGAAAGCTGTAAAGAGAAAGATCCAGCTGTTGTTAAAAAGGAGGAAAAAGTCTT 720
 DB |||||
 QY 661 CAGGGGAAAGCTGTAAAGAGAAAGATCCAGCTGTTGTTAAAAAGGAGGAAAAAGTCTT 720
 DB |||||
 QY 721 CATGAATTTATTTGGGACGGACATTAAGGTGCTGTGCAGAAATGAATCGGCGAGGGAATGC 780
 DB |||||

```
QY 1861 AAATGGGAGTTTCCAAGAGAAAAATTTAGAGTTTGGGAAAGGTACTAGGATCAGGTGCTTTT 1920
DB 1861 AAATGGGAGTTTCCAAGAGAAAAATTTAGAGTTTGGGAAAGGTACTAGGATCAGGTGCTTTT 1920
QY 1921 GGAAGATGATGAACGCAACAGCTTATGGAATAGCAAAACAGGAGTCTCAATCCAGGTT 1980
DB 1921 GGAAGATGATGAACGCAACAGCTTATGGAATAGCAAAACAGGAGTCTCAATCCAGGTT 1980
QY 1981 GCCGTCAAAATGCTGAAGAAAAAGCAGACAGCTCTGAAGAGAGGCACTCATGTGCAGAA 2040
DB 1981 GCCGTCAAAATGCTGAAGAAAAAGCAGACAGCTCTGAAGAGAGGCACTCATGTGCAGAA 2040
QY 2041 CTCAAGATGATGACCCAGCTGGGAGCCACGAGATATTTGTAACCTCTGCGGGCGTGC 2100
DB 2041 CTCAAGATGATGACCCAGCTGGGAGCCACGAGATATTTGTAACCTCTGCGGGCGTGC 2100
QY 2101 ACATGTCAGAACCAATTTACTTGAATTTTGAATCTGCTGATGATGATCTTCTCAAC 2160
DB 2101 ACATGTCAGAACCAATTTACTTGAATTTTGAATCTGCTGATGATGATCTTCTCAAC 2160
QY 2161 TATCTAAGAGTAAAGAGAAAAATTTTCAAGGACTTGGACAGAGATTTTCAAGGAACAC 2220
DB 2161 TATCTAAGAGTAAAGAGAAAAATTTTCAAGGACTTGGACAGAGATTTTCAAGGAACAC 2220
QY 2221 AATTTTCAGTTTACCCCACTTTTCCAAATCACATCCAAATTCAGGATGCTGCTTCAAGA 2280
DB 2221 AATTTTCAGTTTACCCCACTTTTCCAAATCACATCCAAATTCAGGATGCTGCTTCAAGA 2280
QY 2281 GAAGTTTCAGATACACCCGAGCTCGGATCAAAATCTCAGGCTTTCATGGGAATTCATTCCAC 2340
DB 2281 GAAGTTTCAGATACACCCGAGCTCGGATCAAAATCTCAGGCTTTCATGGGAATTCATTCCAC 2340
QY 2341 TCTGAAGATGAATTTGAATATGAAGAACCAAAAGGCTGGAAGAGAGAGGAGACTTGAAT 2400
DB 2341 TCTGAAGATGAATTTGAATATGAAGAACCAAAAGGCTGGAAGAGAGAGGAGACTTGAAT 2400
QY 2401 GTGCTTACATTTGAAGATCTTCTTTGCTTTGCTATCAAGTTTGCACAAAGGAATGGAATTT 2460
DB 2401 GTGCTTACATTTGAAGATCTTCTTTGCTTTGCTATCAAGTTTGCACAAAGGAATGGAATTT 2460
QY 2461 CTGGAATTTAAGTCTGTGTTTCAAGAGACCTGGCGGAGGAACTGCTGTGACCCAC 2520
DB 2461 CTGGAATTTAAGTCTGTGTTTCAAGAGACCTGGCGGAGGAACTGCTGTGACCCAC 2520
QY 2521 GGGAAAGTGTGAAGATATGTGACTTTGGATTTGGCTCGAGATATCATGAGTATTTCCAAAC 2580
DB 2521 GGGAAAGTGTGAAGATATGTGACTTTGGATTTGGCTCGAGATATCATGAGTATTTCCAAAC 2580
QY 2581 TATGTTGTGAGGGCAATGCCCGTCTGCTGTAAAATGAGTGGCCCGCCGAAAGCCTGTTT 2640
DB 2581 TATGTTGTGAGGGCAATGCCCGTCTGCTGTAAAATGAGTGGCCCGCCGAAAGCCTGTTT 2640
QY 2641 GAAGGCATCTACACCATTAAGAGTATGCTGGTCAATATGGAATATTAATCTGTGGGAATC 2700
DB 2641 GAAGGCATCTACACCATTAAGAGTATGCTGGTCAATATGGAATATTAATCTGTGGGAATC 2700
QY 2701 TTCTCATTGTTGTAATCTTACCTGCGATTCGGTTTGAATCTTCACTTACAACTG 2760
DB 2701 TTCTCATTGTTGTAATCTTACCTGCGATTCGGTTTGAATCTTCACTTACAACTG 2760
QY 2761 ATTCAAAATGGAATTTAAAATGATCAGCCATTTTATGCTTACAGAGAAATATACATTATA 2820
DB 2761 ATTCAAAATGGAATTTAAAATGATCAGCCATTTTATGCTTACAGAGAAATATACATTATA 2820
QY 2821 ATGCAATCTGCTGGGCTTTTGAATCTCAAGGAAACGGCCATCTTCCCTAAATTTGACTCG 2880
DB 2821 ATGCAATCTGCTGGGCTTTTGAATCTCAAGGAAACGGCCATCTTCCCTAAATTTGACTCG 2880
QY 2881 TTTTATGATGTCAGCTGGCAGATCGAGAGAGCGATGATCAGATGTCGATGCGCGT 2940
DB 2881 TTTTATGATGTCAGCTGGCAGATCGAGAGAGCGATGATCAGATGTCGATGCGCGT 2940
```

```
QY 2941 GTTTCGGAATGCTCTCACACCTACCAAAACAGGCGACCTTTTCAGCAGAGAGATGATTTG 3000
DB 2941 GTTTCGGAATGCTCTCACACCTACCAAAACAGGCGACCTTTTCAGCAGAGAGATGATTTG 3000
QY 3001 GGGCTACTCTCTCCGAGGCTCAGGTGGAAGATTCGTAGAGGAAACAAATTTAGTTTTAAGG 3060
DB 3001 GGGCTACTCTCTCCGAGGCTCAGGTGGAAGATTCGTAGAGGAAACAAATTTAGTTTTAAGG 3060
QY 3061 ACTTCATCCCTCCACCTATCCCTAACAGGCTGTAGATTTACCAAAACAAAGATTAATTTCA 3120
DB 3061 ACTTCATCCCTCCACCTATCCCTAACAGGCTGTAGATTTACCAAAACAAAGATTAATTTCA 3120
QY 3121 CACTAAAAGAAAAATCTATTAATCAACTGCTTACAGAGCTTTTCTCTAGAGCCGCTCT 3180
DB 3121 CACTAAAAGAAAAATCTATTAATCAACTGCTTACAGAGCTTTTCTCTAGAGCCGCTCT 3180
QY 3181 GCGTTTACTCTGTTTTCAAAGGCACTTTTGTAAAATCAAAATCATCTCTCACAGGCGAG 3240
DB 3181 GCGTTTACTCTGTTTTCAAAGGCACTTTTGTAAAATCAAAATCATCTCTCACAGGCGAG 3240
QY 3241 GAGGAGCTGATAATGAACCTTTTATTTGGAGCAATTTGATCTGCATCCAGGCTTCTCTCAGGCG 3300
DB 3241 GAGGAGCTGATAATGAACCTTTTATTTGGAGCAATTTGATCTGCATCCAGGCTTCTCTCAGGCG 3300
QY 3301 GCTTGAGTGAATTTGCTGACCTGAAAGTACAGTATATTTCTTTGTAATAATACATAAAACAAAAGC 3360
DB 3301 GCTTGAGTGAATTTGCTGACCTGAAAGTACAGTATATTTCTTTGTAATAATACATAAAACAAAAGC 3360
QY 3361 ATTTTGTCTAGGGAAGCTAATATGATTTTTTAAAGTCTATGTTTAAATATATATGTAAG 3420
DB 3361 ATTTTGTCTAGGGAAGCTAATATGATTTTTTAAAGTCTATGTTTAAATATATATGTAAG 3420
QY 3421 TTTTTCAGCTATTTTAGTGATATATTTTATGGGTGGGAATAAAAATTTCTACTACAGA 3476
DB 3421 TTTTTCAGCTATTTTAGTGATATATTTTATGGGTGGGAATAAAAATTTCTACTACAGA 3476
```

RESULT 11

```
AAAT00802
ID AAAT00802 standard; DNA; 3475 Bp.
XX AC AAAT00802;
XX DT 19-MAR-1996 (first entry)
XX DE Human Flk2/flt3 tyrosine kinase receptor gene.
XX KW Tyrosine kinase receptor; flk2; flt3; agonist; monoclonal antibody;
XX KW haematopoiesis; hypoplasia; anaemia; thrombocytopenia; stem cell; ss.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX CDS 58..3039
XX FT /*tag= a
XX FT /product= flk2/flt3_tyrosine_kinase_receptor
XX PN W09527062-A1.
XX XX 12-OCT-1995.
XX PF 23-MAR-1995; 95WO-US03718.
XX PR 04-APR-1994; 9AUS-0222299.
XX PA (GETH ) GENENTECH INC.
XX PI Bennett BD, Broz SD, Matthews W, Zeigler FC;
XX DR WPI; 1995-358636/46.
XX DR P-FSDB; AAR81869.
XX PT Agonist antibody which binds to flk2-flt3 tyrosine kinase receptor
```

PT - enhances proliferation of haematopoietic stem cells, in the
PT treatment of hypoplasia, anaemia, etc.

PS Disclosure; Page 41-44; 59pp; English.

XX DNA (AAT00801) coding for murine stem cell tyrosine kinase receptor flk2
CC (also called flt3) (AAR81869) was obtained by RT-PCR amplification of RNA
CC isolated from mid-gestation mouse foetal livers using primers based
CC on the murine flt3 sequence, and subcloning of the product into pRK5.1.
CC An flk2/flt3 extracellular domain-IgG1 Fc fusion protein was
CC constructed and used to raise agonist antibodies able to bind to, and
CC activate, flk2/flt3. The human flk2/flt3 receptor gene (AAT00802) and
XX amino acid sequence (AAR81869) are also given.

SQ Sequence 3475 BP; 1042 A; 709 C; 784 G; 940 T; 0 other;

Query Match 99.1%; Score 3470.2; DB 16; Length 3475;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3472; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy	1	CGAGCGGCATCCGAGGGCTGGCGCGCGCCCTGGGGACCCCGGCTCCCGAGGCCATG	60
Db	1	CGAGCGGCATCCGAGGGCTGGCGCGCGCCCTGGGGACCCCGGCTCCCGAGGCCATG	60
Qy	61	CCGGCGTTGGCGCGCGACCGCGGCACCGTGGCGCTGCTGTTGTTTCTGCAATGATA	120
Db	61	CCGGCGTTGGCGCGCGACCGCGGCACCGTGGCGCTGCTGTTGTTTCTGCAATGATA	120
Qy	121	TTTGGGACTATTCAAAATCAAGATCTGCTGTGATCAAGTGTTTAAATCAATCAAG	180
Db	121	TTTGGGACTATTCAAAATCAAGATCTGCTGTGATCAAGTGTTTAAATCAATCAAG	180
Qy	181	AACATGATTCATCAGTGGGAGTCATCATATATCCATGGTATCAGATCCCGGAA	240
Db	181	AACATGATTCATCAGTGGGAGTCATCATATATCCATGGTATCAGATCCCGGAA	240
Qy	241	GACCTCGGCTGTGGTTCAGACCCCGAGCTCAGGACAGTGTACGAAGCTGCCCTGTG	300
Db	241	GACCTCGGCTGTGGTTCAGACCCCGAGCTCAGGACAGTGTACGAAGCTGCCCTGTG	300
Qy	301	GAATGGATGATCTGCTTCCATCACA CTGCAAGTGTGTCGATGCCCGGGAACATT	360
Db	301	GAATGGATGATCTGCTTCCATCACA CTGCAAGTGTGTCGATGCCCGGGAACATT	360
Qy	361	TCCTGTCTCGGCTTTAAGCAGAGCTCCCTGAAATTCGCGCCACATTTGATTACAA	420
Db	361	TCCTGTCTCGGCTTTAAGCAGAGCTCCCTGAAATTCGCGCCACATTTGATTACAA	420
Qy	421	AACAGAGAGTGTGTTCCATGTCATTTGAAATGACAGAAACCCAGCTGGAGATAC	480
Db	421	AACAGAGAGTGTGTTCCATGTCATTTGAAATGACAGAAACCCAGCTGGAGATAC	480
Qy	481	CTACTTTTATTCAGAGTGAAGTACCAATATACCAATATTTTACAGTGAATAGA	540
Db	481	CTACTTTTATTCAGAGTGAAGTACCAATATACCAATATTTTACAGTGAATAGA	540
Qy	541	AATACCTGCTTTACACATTAAGAGACCTTACTTTAGAAAAATGGAACACGAGCC	600
Db	541	AATACCTGCTTTACACATTAAGAGACCTTACTTTAGAAAAATGGAACACGAGCC	600
Qy	601	CTGCTGTGATATCTGAGAGGTTCCAGAGCGGATCGTGAATCGGTCTTGGATTCA	660
Db	601	CTGCTGTGATATCTGAGAGGTTCCAGAGCGGATCGTGAATCGGTCTTGGATTCA	660
Qy	661	CAGGGGAAAGCTGTAAAGAGAAAGTCCAGCTGTTGTTAAAGAGGAGAAAAAGTGCTT	720
Db	661	CAGGGGAAAGCTGTAAAGAGAAAGTCCAGCTGTTGTTAAAGAGGAGAAAAAGTGCTT	720
Qy	721	CATGAATTTTGGGACGACATAGGTGCTGTGCCAGAAATGAATGGGACGGAATGC	780
Db	721	CATGAATTTTGGGACGACATAGGTGCTGTGCCAGAAATGAATGGGACGGAATGC	780
Qy	781	ACCAGGCTGTTCAATAGATCTAAATCAAACTCTCAGACCATTCATGCTCCCAATTTAT	840

Db	781	ACCAGGCTGTTCAATAGATCTAAATCAAACTCTCAGACCATTCGCCCAATTTATTT	840
Qy	841	CTTAAAGTAGGGAAACCCCTTATGGATAAGGTCAAAGCTGTTCAATGTAACCAATGATTC	900
Db	841	CTTAAAGTAGGGAAACCCCTTATGGATAAGGTCAAAGCTGTTCAATGTAACCAATGATTC	900
Qy	901	GGGCTCACCTCGGAATTAGAAACAAACACACCTCGAGGGGGCAACTTCTGATGATGAGT	960
Db	901	GGGCTCACCTCGGAATTAGAAACAAACACACCTCGAGGGGGCAACTTCTGATGATGAGT	960
Qy	961	ACCTATTCAACAAACAGAACTATGATAGGATTTCTGTTGCTTTTGTATCATCAGTGGCA	1020
Db	961	ACCTATTCAACAAACAGAACTATGATAGGATTTCTGTTGCTTTTGTATCATCAGTGGCA	1020
Qy	1021	AGAAAGCACCGGATATACACTTTGTTCTCTTCAAAGCATCCAGTCAATCAGCTTTG	1080
Db	1021	AGAAAGCACCGGATATACACTTTGTTCTCTTCAAAGCATCCAGTCAATCAGCTTTG	1080
Qy	1081	GTTACCATCGTAGGAAGGGATTTATTAATGCTACCAATTCAGTGAAGATTTAGAAAT	1140
Db	1081	GTTACCATCGTAGGAAGGGATTTATTAATGCTACCAATTCAGTGAAGATTTAGAAAT	1140
Qy	1141	GACCAATATGAAGATTTTGTGTTTCTGTGAGTAAAGCTACCCACAAATCAGATCT	1200
Db	1141	GACCAATATGAAGATTTTGTGTTTCTGTGAGTAAAGCTACCCACAAATCAGATCT	1200
Qy	1201	ACGTGGACCTTCTCTCGAAATCAATTTCTGTGAGCAAAAGGGTCTTGATACCGATAC	1260
Db	1201	ACGTGGACCTTCTCTCGAAATCAATTTCTGTGAGCAAAAGGGTCTTGATACCGATAC	1260
Qy	1261	AGCATATCAAGTTTTGAATCATTAAGCACAGCCAGGAGAAATATATATCCATGCAGAA	1320
Db	1261	AGCATATCAAGTTTTGAATCATTAAGCACAGCCAGGAGAAATATATATCCATGCAGAA	1320
Qy	1321	AATGATGATGCCAATTTTACCAATATGTTACGCTGAATATAAGAGGAACTCAAGTG	1380
Db	1321	AATGATGATGCCAATTTTACCAATATGTTACGCTGAATATAAGAGGAACTCAAGTG	1380
Qy	1381	CTCGAGAGCATCGGCAAGTCAAGCTGCTGTTTCTCGGATGATACCATTCATCT	1440
Db	1381	CTCGAGAGCATCGGCAAGTCAAGCTGCTGTTTCTCGGATGATACCATTCATCT	1440
Qy	1441	TGGACCTGGAAAGAGTGTTCAGACAAGTCTCCCACTGCA CAGAGAGATCA CAGAGAA	1500
Db	1441	TGGACCTGGAAAGAGTGTTCAGACAAGTCTCCCACTGCA CAGAGAGATCA CAGAGAA	1500
Qy	1501	GTCTGGATGAAGGCTTACAGAAAAGTGTTCGACAGTGGGTGTCAGCAGTACTCTA	1560
Db	1501	GTCTGGATGAAGGCTTACAGAAAAGTGTTCGACAGTGGGTGTCAGCAGTACTCTA	1560
Qy	1561	AACATGATGAAGCATATAAAGGGTTCCTGTCAGTCAAGTCTGTCATACAATTCCTTGGC	1620
Db	1561	AACATGATGAAGCATATAAAGGGTTCCTGTCAGTCAAGTCTGTCATACAATTCCTTGGC	1620
Qy	1621	ACATCTTGTGAGACGATCTTTTAAACTCTCCAGGCCCCCTTCCCTTTTCAACAGACAAC	1680
Db	1621	ACATCTTGTGAGACGATCTTTTAAACTCTCCAGGCCCCCTTCCCTTTTCAACAGACAAC	1680
Qy	1681	ATCTCATTTATGCAACAATTTGTTGTTCTCTCTTCTCATTTGTCGTTTTAACCCCTGTA	1740
Db	1681	ATCTCATTTATGCAACAATTTGTTGTTCTCTCTCTTCTCATTTGTCGTTTTAACCCCTGTA	1740
Qy	1741	ATTTGTCAAGTACAAAAGCAATTTAGGTATGAAAAGCAGCTACAGATGGTACAGTG	1800
Db	1741	ATTTGTCAAGTACAAAAGCAATTTAGGTATGAAAAGCAGCTACAGATGGTACAGTG	1800
Qy	1801	ACCGCTCTCTCAGATAATGATCTTCTAGCTTGTGATTTACAGAGAAATGATATGATCTC	1860
Db	1801	ACCGCTCTCTCAGATAATGATCTTCTAGCTTGTGATTTACAGAGAAATGATATGATCTC	1860
Qy	1861	AAATGGAGTTTCAAGAGAAAATTTAGAGTTTGGGAGGTACTAGGATCAGGTGCTTTT	1920

[illegible]

Qy	3001	GGGCTACTCTCCGAGGCTCAGGTCGAAGATTGCTGAGAGAACAATTTAGTTTAAAGG	3060
Db	3001	GGGCTACTCTCTCCGAGGCTCAGGTCGAAGATTGCTGAGAGAACAATTTAGTTTAAAGG	3060
Qy	3061	ACTTTCATCCCTCCACCTATCCCTTAACAGGCTGTAGATTACCAAAACAAGATTAAATTCAT	3120
Db	3061	ACTTTCATCCCTCCACCTATCCCTTAACAGGCTGTAGATTACCAAAACAAGATTAAATTCAT	3120
Qy	3121	CACATAAAGAAAATCTATTATCAACTGCTGCTTCCACGACTTTTCTCTAGAAGCCGTCT	3180
Db	3121	CACATAAAGAAAATCTATTATCAACTGCTGCTTCCACGACTTTTCTCTAGAAGCCGTCT	3180
Qy	3181	GGCTTTACTCTTGTTTTCAAGGGACTTTTGTAAATCAATCATCTGTCAAGGCGAG	3240
Db	3181	GGCTTTACTCTTGTTTTCAAGGGACTTTTGTAAATCAATCATCTGTCAAGGCGAG	3240
Qy	3241	GAGGAGCTGATAATGAATTTATTTGGAGCAATTCATCTGCATCCAAGGCCCTTCTCAGGCCG	3300
Db	3241	GAGGAGCTGATAATGAATTTATTTGGAGCAATTCATCTGCATCCAAGGCCCTTCTCAGGCCG	3300
Qy	3301	GCTTGAAGTGAATTTGTGCTGCTGAAGTACAGTATATCTTGTAAATACATAAAACAAAAGC	3360
Db	3301	GCTTGAAGTGAATTTGTGCTGCTGAAGTACAGTATATCTTGTAAATACATAAAACAAAAGC	3360
Qy	3361	ATTTTGTGAAGGAGAGCTAATATGATTTTAAAGTCTATGTTTAAATAATATGTAAA	3420
Db	3361	ATTTTGTGAAGGAGAGCTAATATGATTTTAAAGTCTATGTTTAAATAATATGTAAA	3420
Qy	3421	TTTTTTCAGCTATTTAGTGATATATTTTATGGGTGGGAATAAAATTTCTACTACAG	3475
Db	3421	TTTTTTCAGCTATTTAGTGATATATTTTATGGGTGGGAATAAAATTTCTACTACAG	3475
RESULT 12			
AAD42484			
ID	AAD42484 standard; cDNA; 3489 BP.		
XX	AAD42484;		
XX	15-NOV-2002 (first entry)		
DT	Human receptor protein tyrosine kinase, flk-2 cDNA.		
DE	Human; receptor protein tyrosine kinase; pTK; haematopoietic cell; growth factor; therapeutic; macrocytic anaemia; aplastic anaemia; bone marrow damage; cancer; chemotherapy; radiation; osteopathic; flk-2; gene; ss.		
XX	Homo sapiens.		
OS	Key		
PH	CDS		
FT	Location/Qualifiers		
FT	58..3039		
FT	/*tag= a		
FT	/product= "Human flk-2 protein"		
FT	58..138		
FT	/*tag= b		
FT	139..3036		
FT	/*tag= c		
FT	/product= "Human mature flk-2 protein"		
FT	139..1689		
FT	/*tag= d		
FT	/note= "Encodes extracellular domain"		
FT	1690..1746		
FT	/*tag= e		
FT	/note= "Encodes transmembrane domain"		
FT	1747..3036		
FT	/*tag= f		
FT	/note= "Encodes intracellular domain"		
XX	US2002072077-A1.		
PW	13-JUN-2002.		
XX			

XX 31-JUL-2001; 2001US-0919408.
PF 19-NOV-1992; 92US-0977451.
PR 30-APR-1993; 93US-0055269.
PR 31-OCT-1994; 94US-0252498.
PR 15-FEB-1996; 96US-0601891.
PR 10-FEB-1998; 98US-0021324.
PR 10-DEC-1998; 98US-0208786.
PR 02-APR-1991; 91US-0679666.
PR 28-JUN-1991; 91US-0728913.
PR 15-NOV-1991; 91US-0793065.
PR 24-DEC-1991; 91US-0813593.
XX (LEMI/) LEMISCHKA I R.
PA Lemischka IR;
XX WPI; 2002-607237/65.
DR P-PSDB; AAE25819.
XX New protein tyrosine kinase expressed in primitive hematopoietic cells
PT (HC) and not expressed in mature HC, and ligands for the protein, for
PT stimulating proliferation of primitive hematopoietic stem cells
XX Claim 9; Page 22-25; 64pp; English.
XX The present invention relates to receptor protein tyrosine kinase (pTK)
CC expressed in primitive haematopoietic cells and not expressed in mature
CC haematopoietic cells, polynucleotides encoding such proteins and ligands
CC for the protein. Ligands which bind to pTK of the invention are useful
CC for stimulating the proliferation and/or differentiation of primitive
CC mammalian haematopoietic stem cells. The receptor pTK molecules are
CC useful for stimulating the self renewal of the totipotent haematopoietic
CC stem cell and to stimulate the development of all cells of haematopoietic
CC system both in vitro and in vivo. The ligands for the receptors act as
CC haematopoietic growth factors. The ability of the ligands to stimulate
CC proliferation of stem cells both in vitro and in vivo has important
CC therapeutic applications such as treating humans whose primitive stem
CC cells do not sufficiently undergo self-renewal. It is also useful in
CC conditions that occur when defects in haematopoietic stem cells or their
CC related growth factors depress the number of white blood cells such as
CC macrocytic and aplastic anaemia or bone marrow damage resulting from
CC cancer chemotherapy and radiation. The present sequence is a cDNA
XX encoding human receptor pTK, fik-2.
SQ Sequence 3489 BP; 1066 A; 707 C; 780 G; 936 T; 0 other;

Query Match
Best Local Similarity 97.6%; Score 3417; DB 24; Length 3489;
Matches 3489; Conservative 0; Mismatches 0; Indels 12; Gaps 6;

Qy 1 CGAGCGGCATCCGAGGGCTGGGCGGCGCCCTGGGGGACCCCGGCTCCGAGGCCATG 60
Db 1 CGAGCGGCATCCGAGGGCTGGGCGGCGCCCTGGGGGACCCCGGCTCCGAGGCCATG 60
Qy 61 CCGGCGTTGGCGCGGACCGGCGACCGGTCGCTGCTGTTTTCGCAATGATA 120
Db 61 CCGGCGTTGGCGCGGACCGGCGACCGGTCGCTGCTGTTTTCGCAATGATA 120
Qy 121 TTGCGGATTTACAAATCAAGATCTGCTGATCAAGTGTGTTTAAATCAATCAAG 180
Db 121 TTGCGGATTTACAAATCAAGATCTGCTGATCAAGTGTGTTTAAATCAATCAAG 180
Qy 181 AACAAATGATTCATAGTGGGGAAGTCATCATCATATCCATGGTATCAGATCCCGGAA 240
Db 181 AACAAATGATTCATAGTGGGGAAGTCATCATCATATCCATGGTATCAGATCCCGGAA 240
Qy 241 GACCTCGGGTGTGCTTGAGACCCGAGAGCTCAGGACAGTGTACCAAGCTGCCGCTGTG 300
Db 241 GACCTCGGGTGTGCTTGAGACCCGAGAGCTCAGGACAGTGTACCAAGCTGCCGCTGTG 300
Qy 301 GAAGTGGATGTATCTGCTTCCATCACTGCAAGTGTGCTGATGCCCGGGAACATT 360

Db 301 GAAGTGGATGTATCTGCTTCCATCACTGCAAGTGTGCTGATGCCCGGGAACATT 360
Qy 361 TCTGTCCTGGGTCTTTAAGACACAGCTCCCTGGAATTGCGAGCCACATTTTGAATTCAA 420
Db 361 TCTGTCCTGGGTCTTTAAGACACAGCTCCCTGGAATTGCGAGCCACATTTTGAATTCAA 420
Qy 421 AACAGAGGAGTGTGTTTCCATGTCATTTTGAATAACACAGAAACCCAGCTGGAGATAC 480
Db 421 AACAGAGGAGTGTGTTTCCATGTCATTTTGAATAACACAGAAACCCAGCTGGAGATAC 480
Qy 481 CTACTTTTATTCAGAGTGAAGCTACCAATTTACACAATATTGTTTACAGTGTATAGA 540
Db 481 CTACTTTTATTCAGAGTGAAGCTACCAATTTACACAATATTGTTTACAGTGTATAGA 540
Qy 541 AATAACCTCTCTTACACATTAAAGAGACCTTACTTTTAGAAAAATGAAAAACAGGAGCC 600
Db 541 AATAACCTCTCTTACACATTAAAGAGACCTTACTTTTAGAAAAATGAAAAACAGGAGCC 600
Qy 601 CTGCTCTCATATCTGAGAGCGTTCCAGAGCCGATCGTGAATGGGTGCTTTCGATTC 660
Db 601 CTGCTCTCATATCTGAGAGCGTTCCAGAGCCGATCGTGAATGGGTGCTTTCGATTC 660
Qy 661 CAGGGGGAAGCTGTAAAGAGAAAGTCCAGCTGTTTGTAAAGAGGAGAAAGTGCCT 720
Db 661 CAGGGGGAAGCTGTAAAGAGAAAGTCCAGCTGTTTGTAAAGAGGAGAAAGTGCCT 720
Qy 721 CATGAATTTATTTGGGACGGACATAAAGTCTCTGTCGAGAAATGAACCTGGGAGGAAATGC 780
Db 721 CATGAATTTATTTGGGACGGACATAAAGTCTCTGTCGAGAAATGAACCTGGGAGGAAATGC 780
Qy 781 ACCAGCTGTTTCAATATAGATCTAAATCAAACTCTCTCAGACACATTTGATTTATTT 840
Db 781 ACCAGCTGTTTCAATATAGATCTAAATCAAACTCTCTCAGACACATTTGATTTATTT 840
Qy 841 CTTAAAGTAGGGGACCCCTTATGATAGGTGCAAGCTGTTTCAATGTAACCATGATTC 900
Db 841 CTTAAAGTAGGGGACCCCTTATGATAGGTGCAAGCTGTTTCAATGTAACCATGATTC 900
Qy 901 GGGCTCACCTGGGAATTAGAAAAAAGACCTCGAGGAGGCACTACTTTTCAGATGAGT 960
Db 901 GGGCTCACCTGGGAATTAGAAAAAAGACCTCGAGGAGGCACTACTTTTCAGATGAGT 960
Qy 961 ACCTATTCAACAAACAGAACTATGATACGGAATCTGTTTGTATCATCAGTGGCA 1020
Db 961 ACCTATTCAACAAACAGAACTATGATACGGAATCTGTTTGTATCATCAGTGGCA 1020
Qy 1021 AGAAACGACCGGATACACTGTTCTCTTCAAGAGCATCCAGTCAATCAGCTTTG 1080
Db 1021 AGAAACGACCGGATACACTGTTCTCTTCAAGAGCATCCAGTCAATCAGCTTTG 1080
Qy 1081 GTTACCATCTAGGAAAGGGATTTATAATGCTACCAATTCAGTGAAGATTTATGAATTT 1140
Db 1081 GTTACCATCTAGGAAAGGGATTTATAATGCTACCAATTCAGTGAAGATTTATGAATTT 1140
Qy 1141 GACCAATATGAAGAGTGTGTTTCTGTCAGGTTTAAAGCCCTACCCAAATCAGATCT 1200
Db 1141 GACCAATATGAAGAGTGTGTTTCTGTCAGGTTTAAAGCCCTACCCAAATCAGATCT 1200
Qy 1201 ACGTGGACCTTCTCTCGAAATCAATTCCTTGTGAGCAAAAAGGGTCTTTGATAACGATAC 1260
Db 1201 ACGTGGACCTTCTCTCGAAATCAATTCCTTGTGAGCAAAAAGGGTCTTTGATAACGATAC 1260
Qy 1261 AGCATATCCAAAGTGTGCAATCAATAGCACAGCCAGGAGGATATATATTCATCCAGAA 1320
Db 1261 AGCATATCCAAAGTGTGCAATCAATAGCACAGCCAGGAGGATATATATTCATCCAGAA 1320
Qy 1321 AATGATGATGCCCAATTTTACCAAAATGTTTACGCTGGAATATAAGAAAGAAACCTCAAGTG 1380
Db 1321 AATGATGATGCCCAATTTTACCAAAATGTTTACGCTGGAATATAAGAAAGAAACCTCAAGTG 1380
Qy 1381 CTCGAGAAAGCATCGGCAAGTCAAGCGTCTGTTTCTCGGATGGATACCCATTCATCT 1440

Db 1381 CTCGAGAAAGCAATCGGCAAGTCAGGGCTCCTGTTCTTCGGATGGGATACCCATTACCATCT 1440
Qy 1441 TGGACCTGGAAGAAAGTGTTCAGACAAGTCTCCAACTGCAAGAGAGATCAAGAGGA 1500
Db 1441 TGGACCTGGAAGAAAGTGTTCAGACAAGTCTCCAACTGCAAGAGAGATCAAGAGGA 1500
Qy 1501 GTCTGGAATAGAAAGCGCTAAACAGAAAGTGTGTGGACAGTGGGTGTCGAGCAGTACTCTA 1560
Db 1501 GTCTGGAATAGAAAGCGCTAAACAGAAAGTGTGTGGACAGTGGGTGTCGAGCAGTACTCTA 1560
Qy 1561 AACATGAGTGAAGCCATTAAGAGGTTCTGTGTCAGTGTGTGTCATACAAATCCCTGTGC 1620
Db 1561 AACATGAGTGAAGCCATTAAGAGGTTCTGTGTCAGTGTGTGTCATACAAATCCCTGTGC 1620
Qy 1621 ACATCTGTGACAGCATCTTTTAAACTCTCAGGCCCTTCCCTTTTCATCAAGACACAC 1680
Db 1621 ACATCTGTGACAGCATCTTTTAAACTCTCAGGCCCTTCCCTTTTCATCAAGACACAC 1680
Qy 1681 ATCTCATCTATGCAACAAATGGTGTGTGTCTCTCTTCATTTGTCGTTTAAACCTGTCTA 1740
Db 1681 ATCTCATCTATGCAACAAATGGTGTGTGTCTCTCTCTTCATTTGTCGTTTAAACCTGTCTA 1740
Qy 1741 ATTTGTCACAGTACAAAGCAATTTAGGTATGAAGCCAGCTACAGATGGTACAGGTG 1800
Db 1741 ATTTGTCACAGTACAAAGCAATTTAGGTATGAAGCCAGCTACAGATGGTACAGGTG 1800
Qy 1801 ACCGGCTCCTCAGATAATGAGTACTTCTACGTGTGATGAAAGCCAGCTACAGATGATCTC 1860
Db 1801 ACCGGCTCCTCAGATAATGAGTACTTCTACGTGTGATGAAAGCCAGCTACAGATGATCTC 1860
Qy 1861 AAATCGGAGTTTCCAGAGAAAATTTAGAGTTTGGGAAGGTACTAGGATCAGGTGCTTTT 1920
Db 1861 AAATCGGAGTTTCCAGAGAAAATTTAGAGTTTGGGAAGGTACTAGGATCAGGTGCTTTT 1920
Qy 1921 GGAAGAGTGAAGCGCAACAGCTTATGGAATTTAGCAAGAGAGGACCTCATGTCAAGAA 2040
Db 1921 GGAAGAGTGAAGCGCAACAGCTTATGGAATTTAGCAAGAGAGGACCTCATGTCAAGAA 2040
Qy 2041 CTCAGATGATGACCCAGCTGGGAAGCCAGAGAAATTTGTGAACCTGTGGGGCGTGC 2100
Db 2041 CTCAGATGATGACCCAGCTGGGAAGCCAGAGAAATTTGTGAACCTGTGGGGCGTGC 2100
Qy 2101 ACACGTGTCAGGACCAATTTACTGTTTGTGATGATGATGATGATGATGATGATGATGATG 2160
Db 2101 ACACGTGTCAGGACCAATTTACTGTTTGTGATGATGATGATGATGATGATGATGATGATG 2160
Qy 2161 TATCTAAGAGTAAAGAGAAAATTTACAGGACTTGGACAGAGATTTTCAAGGAACAC 2220
Db 2161 TATCTAAGAGTAAAGAGAAAATTTACAGGACTTGGACAGAGATTTTCAAGGAACAC 2220
Qy 2221 AATTTTCAGTTTACCCCACTTTTCAATCAGATCCTCAAAATTTCCAGCATGCTGGTTCAAGA 2280
Db 2221 AATTTTCAGTTTACCCCACTTTTCAATCAGATCCTCAAAATTTCCAGCATGCTGGTTCAAGA 2280
Qy 2281 GAAGTTTCAGATACACCCGAGCTGGATCAAAATCTCAGGGCTTCATGGGAATTCATTTTAC 2340
Db 2281 GAAGTTTCAGATACACCCGAGCTGGATCAAAATCTCAGGGCTTCATGGGAATTCATTTTAC 2340
Qy 2341 TCTGAAGTGAATTTGAATATGAATATGAATATGAATATGAATATGAATATGAATATGAAT 2400
Db 2341 TCTGAAGTGAATTTGAATATGAATATGAATATGAATATGAATATGAATATGAATATGAAT 2400
Qy 2401 GTGCTTACATTTGAAGATCTTCTTCTGTTGATATCAAGTTTCCAAAGGAATGGAATTT 2460
Db 2401 GTGCTTACATTTGAAGATCTTCTTCTGTTGATATCAAGTTTCCAAAGGAATGGAATTT 2460
Qy 2461 CTGGAATTTAAGTGTGTGTTCACAGACCTGGCCGACGAGAACTGTGTGTGTGTGTGTGTGT 2520
Db 2461 CTGGAATTTAAGTGTGTGTTCACAGACCTGGCCGACGAGAACTGTGTGTGTGTGTGTGTGT 2520

RESULT 13
AAQ49756
ID AAQ49756 standard; DNA; 3120 BP.

Qy 2521 GGGAAAGTGGTGAAGATATGTGACTTTGGATTGGCTCGAGATATCATGTGATTTCCAAAC 2580
Db 2521 GGGAAAGTGGTGAAGATATGTGACTTTGGATTGGCTCGAGATATCATGTGATTTCCAAAC 2580
Qy 2581 TATGTTGTTCAGGGCAATGCGCTGCTGTCTGTAAATGAGTGGCCCGGAAAGCCCTGTTT 2640
Db 2581 TATGTTGTTCAGGGCAATGCGCTGCTGTCTGTAAATGAGTGGCCCGGAAAGCCCTGTTT 2640
Qy 2641 GAAGGCATCTACACCATTTAAGAGTGTCTGTCTGTATGGAATATTTACTGTGGAAATC 2700
Db 2641 GAAGGCATCTACACCATTTAAGAGTGTCTGTCTGTATGGAATATTTACTGTGGAAATC 2700
Qy 2701 TTCTCATTGTGTGAATTCCTTACCTGGCATTCCTGGTGTGATGCTAACTTCTACAAACTG 2760
Db 2701 TTCTCATTGTGTGAATTCCTTACCTGGCATTCCTGGTGTGATGCTAACTTCTACAAACTG 2760
Qy 2761 ATTCAAAATGATTTAAATGATCAGCCATTTTATGCTACAGAAATATATACATTATA 2820
Db 2761 ATTCAAAATGATTTAAATGATCAGCCATTTTATGCTACAGAAATATATACATTATA 2820
Qy 2821 ATGCAATCTGCTGGGCTTTTGAATCAAGGAAACCGCCATCCTTCCCTAAATTTGACTCG 2880
Db 2821 ATGCAATCTGCTGGGCTTTTGAATCAAGGAAACCGCCATCCTTCCCTAAATTTGACTCG 2880
Qy 2881 TTTTTPAGGATGTACGTGGCAGATCAGAAAGCGATGTATCAGAAATGTGATGGCCGT 2940
Db 2881 TTTTTPAGGATGTACGTGGCAGATCAGAAAGCGATGTATCAGAAATGTGATGGCCGT 2940
Qy 2941 GTTTCGGAAATGCTTCACACCTTACCAACAGCGGACCTTTTCAGCAGAGATGGAATTTG 3000
Db 2941 GTTTCGGAAATGCTTCACACCTTACCAACAGCGGACCTTTTCAGCAGAGATGGAATTTG 3000
Qy 3001 GGGCTACTCTCTCCGAGGCTCAGGTGCAAGATTTCTAGAGGAAACAAATTTAGTTTAAAG 3060
Db 3001 GGGCTACTCTCTCCGAGGCTCAGGTGCAAGATTTCTAGAGGAAACAAATTTAGTTTAAAG 3060
Qy 3061 ACTTCATCCCTCCACTTATCCCTAACAGGCTGTAGATTACCAAAACAAGATTTTTCAT 3120
Db 3061 ACTTCATCCCTCCACTTATCCCTAACAGGCTGTAGATTACCAAAACAAGATTTTTCAT 3120
Qy 3121 CACTAAAAGAAATCTATATCACTGCTGCTTACAGACCTTTTCTAGAGCCGTCT 3180
Db 3121 CACTAAAAGAAATCTATATCACTGCTGCTTACAGACCTTTTCTAGAGCCGTCT 3180
Qy 3181 GGGTTTACTCTTGTGTTTCAAGGGACTTTTGTAAATCAAAATCAATCCTCTCAAGGCGAG 3240
Db 3181 GGGTTTACTCTTGTGTTTCAAGGGACTTTTGTAAATCAAAATCAATCCTCTCAAGGCGAG 3240
Qy 3179 GGGTTTACTCTTGTGTTTCAAGGG--TTTGTAAATCAAAATCAATCCTCTCAAGGCGAG 3236
Db 3179 GGGTTTACTCTTGTGTTTCAAGGG--TTTGTAAATCAAAATCAATCCTCTCAAGGCGAG 3236
Qy 3241 GAGGAGCTGATAATGAACCTTTATTTGGAGCATTTGATCTGCATCCAGGCCCTTCTCAGGCCG 3300
Db 3241 GAGGAGCTGATAATGAACCTTTATTTAGTGTGATGATGATGATGATGATGATGATGATGAT 3300
Qy 3301 GCTTGAAGTGAATGTGTACCTGAAGTACAGTATATTTCTGTAAATACATAAACAAGC 3360
Db 3301 GCTTGAAGTGAATGTGTACCTGAAGTACAGTATATTTCTGTAAATACATAAACAAGC 3360
Qy 3352 GCTTGAAGTGAATGTGTACCTGAA--ACAGTATATTTTGTAAATACATAAACAAGC 3352
Db 3352 GCTTGAAGTGAATGTGTACCTGAA--ACAGTATATTTTGTAAATACATAAACAAGC 3352
Qy 3361 ATTTTGTCTAAGGAGAGCTTAATATGATTTTAACTGATGTTTAAATATATATGTA 3420
Db 3361 ATTTTGTCTAAGGAGAGCTTAATATGATTTTAACTGATGTTTAAATATATATGTA 3420
Qy 3353 ATTTTGTCTAAGGAGAGCTTAATAT--TTTGTAACTGATGTTTAAATATATATGTA 3410
Db 3353 ATTTTGTCTAAGGAGAGCTTAATAT--TTTGTAACTGATGTTTAAATATATATGTA 3410
Qy 3421 TTTTTCAGCTATTTAGTGTATATTTTATGTTGGTGGGAATATAAATTTCTACTACAGAAAA 3480
Db 3421 TTTTTCAGCTATTTAGTGTATATA--TTATGGTGGGAATATAAATTTCTACTACAGAAAA 3480
Qy 3481 AAAAAAAAAAAAAAAAAAAAAA 3501
Db 3469 AAAAAAAAAAAAAAAAAAAAAA 3489

XX	AAQ49756;
AC	XX
XX	25-MAR-2003 (updated)
DT	10-MAR-1994 (first entry)
DE	XX
XX	ptk gene LptK25.
XX	
KW	ptk; protein tyrosine kinase; catalytic domain; c-kit; megakaryocyte;
KW	lymphocyte; amplification; primer; polymerase chain reaction; PCR; ds.
OS	Homo sapiens.
XX	XX
XX	WO9315201-A1.
PX	XX
PD	05-AUG-1993.
XX	XX
PF	22-JAN-1993; 93WO-US00586.
XX	XX
PR	22-JAN-1992; 92US-0826935.
XX	XX
PA	(NEWE-) NEW ENGLAND DEACONESS HOSPITAL.
XX	XX
PI	Avraham H, Cowley S, Groopman J, Scadden D;
XX	XX
DR	WPI; 1993-320330/40.
XX	XX
PT	New protein tyrosine kinase genes and proteins encoded by genes -
PT	are of human mega-karyocytic origin
XX	XX
PS	Claim 2; Fig 7; 60pp; English.
XX	XX
CC	ptk genes were identified using two sets of degenerative
CC	oligonucleotide primers: a first set which amplifies all ptk DNA
CC	segments (AAQ49743-44) and a second set which amplifies highly
CC	conserved sequences present in the catalytic domain of the c-kit
CC	subgroup of ptk's (AAQ49745-46). The ptk genes identified are described
CC	in AAQ49747-57 and AAR41897-02.
CC	The Lptks are expressed in lymphocytic cells, as well as
CC	megakaryocytic cells.
CC	(Updated on 25-MAR-2003 to correct PN field.)
XX	XX
SQ	Sequence 3120 BP; 945 A; 643 C; 699 G; 833 T; 0 other;
	Query Match 88.2%; Score 3086.4; DB 14; Length 3120;
	Best Local Similarity 99.3%; Pred. No. 0;
	Matches 3099; Conservative 0; Mismatches 21; Indels 0; Gaps 0;
QY	58 ATGCCGCGTTGGCGGCAGCGCGCACCGCTCGTGTCTGTTTTCTGCAATG 117
Dd	1 ATGAGAGCGTTGGCGGCAGCGCGCGCAGCTGCTGTTTTCTGCAATG 60
QY	118 ATATTTCGGACTATTACAATCAAGATCTGCTGTGATCAAGTGTTTTAATCAATCAT 177
Dd	61 ATATTTCGGACTATTACAATCAAGATCTGCTGTGATCAAGTGTTTTAATCAATCAT 120
QY	178 AAGAACAATGATTTCATCAGTGGGGAAGTCATCATATCCATGGTATCAGATCCCCG 237
Dd	121 AAGAACAATGATTTCATCAGTGGGGAAGTCATCATATCCATGGTATCAGATCCCCG 180
QY	238 GAAGACCTCGGTGTGCTTTAGACCACAGAGCTCAGGACAGTGTCGCGCT 297
Dd	181 GAAGACCTCGGTGTGCTTTAGACCACAGAGCTCAGGACAGTGTCGCGCT 240
QY	298 GTGGAAGTGTATGTATCTGCTTCATCACA CTGCAAGTGTGTCGATGCCAGGGAAC 357
Dd	241 GTGGAAGTGTATGTATCTGCTTCATCACA CTGCAAGTGTGTCGATGCCAGGGAAC 300
QY	358 ATTTCCTGCTCTGGGTCITTTAAGCACAGCTCCCTGAATTGCCAGGCCACATTTTGATTTA 417
Dd	301 ATTTCCTGCTCTGGGTCITTTAAGCACAGCTCCCTGAATTGCCAGGCCACATTTTGATTTA 360
QY	418 CAATAACAGAGGAGTGTGTTCCATGGTGTCATNTTTGAAAATGACAGAAACCCAGCTGGAGAA 477

1441	DB	GGAGTCTGGAATAGAAAGCTTAACAGAAAAAGTTGTTGGACAGTGGGTGTCGAGCAGTACT	1500
1558	QY	CTAAACATGAGTGAAGCCATAAAAGGGTTCTCTGGTCAAGTGTGTGTCATACAATTCCTTT	1617
1501	DB	CTAACAATGAGTGAAGCCATAAAAGGGTTCTCTGGTCAAGTGTGTGTCATACAATTCCTTT	1560
1618	QY	GGCACAATCTTTGTGAGACGATCTCTTTTAACTCTCAGGCCCCCTTCCCTTTTATCCAGAC	1677
1561	DB	GGCACAATCTTTGTGAGACGATCTCTTTTAACTCTCAGGCCCCCTTCCCTTTTATCCAGAC	1620
1678	QY	ACATCTCAATCTATGCAACAATTCGGTGTTCCTCTCTCATTTGTCGTTTAAACCCCTG	1737
1621	DB	AACATCTCAATCTATGCAACAATTCGGTGTTCCTCTCTCATTTGTCGTTTAAACCCCTG	1680
1738	QY	CTAATTTGTACAAAGTACAAAAGCAATTTAGGTATGAAAGCCAGCTACAGATGGTACAG	1797
1681	DB	CTAATTTGTACAAAGTACAAAAGCAATTTAGGTATGAAAGCCAGCTACAGATGGTACAG	1740
1798	QY	GTGACCGGTCTCTCAGATAATGAGTACTTCTCTACGTTGATTTTCAGAGAAATATGAATATGAT	1857
1741	DB	GTGACCGGTCTCTCAGATAATGAGTACTTCTCTACGTTGATTTTCAGAGAAATATGAATATGAT	1800
1858	QY	CTCAATGGAGTTTCCAAAGAGAAAATTTAGAGTTTGGGAAAGTACTAGAGTACAGTGCT	1917
1801	DB	CTCAATGGAGTTTCCAAAGAGAAAATTTAGAGTTTGGGAAAGTACTAGAGTACAGTGCT	1860
1918	QY	TTTGGAAAAAGTGATGAACGCAACAGCTTATGGAATTTAGCAAAACAGGAGTCTCAATCCAG	1977
1861	DB	TTTGGAAAAAGTGATGAACGCAACAGCTTATGGAATTTAGCAAAACAGGAGTCTCAATCCAG	1920
1978	QY	GTTCGCGTCAAAATGCTGAAAAGAAAAAGCAGACGCTCTGAAAGAGAGGCATCATGTCA	2037
1921	DB	GTTCACCGTCAAAATGCTGAAAAGAAAAAGCAGACGCTCTGAAAGAGAGGCATCATGTCA	1980
2038	QY	GAACTCAAGATGATGACCCAGCTGGGAGCCACGAGAAATTTGCTGAACTGCTGGGGGG	2097
1981	DB	GAACTCAAGATGATGACCCAGCTGGGAGCCACGAGAAATTTGCTGAACTGCTGGGGGG	2040
2098	QY	TGCACACTGTCCAGGACCAATTTACTTTGATTTTTGGAATACTGTGCTATGGTGATCTTCTC	2157
2041	DB	TGCACACTGTCCAGGACCAATTTACTTTGATTTTTGGAATACTGTGCTATGGTGATCTTCTC	2100
2158	QY	AACATCTTAAGAAAGTAAAGAGAAAAATTTTCACGAGCTTGGACAGAGATTTTCAAGGAA	2217
2101	DB	AACATCTTAAGAAAGTAAAGAGAAAAATTTTCACGAGCTTGGACAGAGATTTTCAAGGAA	2160
2218	QY	CACAAATTCTAGTTTATCCCACTTTCCAAATCATCTCCAAATTCAGAGTATGCTGGTTCA	2277
2161	DB	CACAAATTCTAGTTTATCCCACTTTCCAAATCATCTCCAAATTCAGAGTATGCTGGTTCA	2220
2278	QY	AGAGAGTTCAGATACACCCGGACTCGGATCAAACTCTCAGGGCTTCATGGGAATTCATTT	2337
2221	DB	AGAGAGTTCAGATACACCCGGACTCGGATCAAACTCTCAGGGCTTCATGGGAATTCATTT	2280
2338	QY	CACCTCTGAAGATGAAATTTGAATATGAAAACCAAAAAAGGCTGGAAGAGAGGAGGACTTG	2397
2281	DB	CACCTCTGAAGATGAAATTTGAATATGAAAACCAAAAAAGGCTGGAAGAGAGGAGGACTTG	2340
2398	QY	AATGTGCTTACATTTGGAAGATCTTCTTTGCTTTGTCATATCAAGTTGCCAAGGAATGGAA	2457
2341	DB	AATGTGCTTACATTTGGAAGATCTTCTTTGCTTTGTCATATCAAGTTGCCAAGGAATGGAA	2400
2458	QY	TTTTCTGGAATTTAAGTCTGTGTTTCACAGAGACTGGCCGCCAGGAACGCTGTGTCACC	2517
2401	DB	TTTTCTGGAATTTAAGTCTGTGTTTCACAGAGACTGGCCGCCAGGAACGCTGTGTCACC	2460
2518	QY	CACGGGAAAGTGGTGAAGATATGTGACTTTGGATTTGGCTCGAGATATCATGAGTGATTC	2577
2461	DB	CACGGGAAAGTGGTGAAGATATGTGACTTTGGATTTGGCTCGAGATATCATGAGTGATTC	2520
2578	QY	AACATATGTTGTACGGGCAATGCCCCGTCTGCTCTTAAATAGGATGGCCCCCGAAGGCTG	2637
2521	DB	AACATATGTTGTACGGGCAATGCCCCGTCTGCTCTTAAATAGGATGGCCCCCGAAGGCTG	2580

Qy	2638	TTTGAAGGCATCTACACCATTAAGAGTGATCTGTGTCATATGGAATATTACTGTGGAA	2699
Db	2581	TTTGAAGGCATCTACACCATTAAGAGTGATCTGTGTCATATGGAATATTACTGTGGAA	2640
Qy	2698	ATCTTCTCACTCGGTGTGAATCCTTACCCTGGCAATCCGGTTGATGCTAACTCTTACAAA	2757
Db	2641	ATCTTCTCACTCGGTGTGAATCCTTACCCTGGCAATCCGGTTGATGCTAACTCTTACAAA	2700
Qy	2758	CTGATTCMAAATCGATTTAAATTTGGATCAGCCATTTTATGCTACAGAGAATAATACATT	2817
Db	2701	CTGATTCMAAATCGATTTAAATTTGGATCAGCCATTTTATGCTACAGAGAATAATACATT	2760
Qy	2818	ATAATGCAATCCTGCTGGGCTTTTGACTCAAGGAAACGGCCATCTTCCCTAAATTTGACT	2877
Db	2761	ATAATGCAATCCTGCTGGGCTTTTGACTCAAGGAAACGGCCATCTTCCCTAAATTTGACT	2820
Qy	2878	TCGTTTTTAGATGTCAGCTGGCAGATGTCAGAGAGCGGATGTATCAGATGTGGATGGC	2937
Db	2821	TCGTTTTTAGATGTCAGCTGGCAGATGTCAGAGAGCGGATGTATCAGATGTGGATGGC	2880
Qy	2938	CGTGTTCGGAATGCTCCTCACACCTACCAAAAACAGGCGACCTTTTCAGCAGAGATGGAT	2997
Db	2881	CGTGTTCGGAATGCTCCTCACACCTACCAAAAACAGGCGACCTTTTCAGCAGAGATGGAT	2940
Qy	2998	TTGGGGCTACTCTCTCCGCGAGGCTCAGGTCGAAGATTCGTAGAGGACAATTTAGTTTTA	3057
Db	2941	TTGGGGCTACTCTCTCCGCGAGGCTCAGGTCGAAGATTCGTAGAGGACAATTTAGTTTTA	3000
Qy	3058	AGGACTTCATCCCTCCACCTATCCTCAACAGGCTGTAGATTACCAAAACAAGATTAAATTT	3117
Db	3001	AGGACTTCATCCCTCCACCTATCCTCAACAGGCTGTAGATTACCAAAACAAGATTAAATTT	3060
Qy	3118	CATCACTAAAGAAAATCTATTATCAACTCTGCTGCTTCCACGAGCTTTTCTCTAGAACGG	3177
Db	3061	CATCACTAAAGAAAATCTATTATCAACTCTGCTGCTTCCACGAGCTTTTCTCTAGAACGG	3120

RESULT 14

AA03096

ID AAT03096 standard; DNA; 3120 BP.

AA
AC
AAT03096:XX
AC
HH10050,

DT 14-FEB-19

5
4
0
7
4
4
4
X

Protein t

XX

KW Protein t

KW different

XX

OS Homo sapi

XX

PN WO9527061

XX

PD 12-OCT-19

XX

PF 04-APR-15

[illegible]

PR 04-APR-15

XX (1957)

PA (GETH) C

XX
PT
Bennett B

PI. Bennett
PI. Wood wt.

XX
TJ
NOON;

WPT. 1995

XX
XX
WET, TSSC

PT
Asonist a

PT kinase(s)

PT domain are

PT modulation


```
Db 2221 TCACATCCAAATTCAGCAGATGCTGCTTCAAGAGAAAGTTTCAGATACACCCGGACTCGGAT 2280
Qy 2308 CAAATCTCAGGCTTTCATGGGAATTCATTTCACTCTGAAGATGAAATTCGAATATGAAAAC 2367
Db 2281 CAAATCTCAGGCTTTCATGGGAATTCATTTCACTCTGAAGATGAAATTCGAATATGAAAAC 2340
Qy 2368 CAAAAAGGCTGGAAGAGAGAGAGACTTGAATGCTTACATTTGAAGATCTTCTTTGC 2427
Db 2341 CAAAAAGGCTGGAAGAGAGAGAGACTTGAATGCTTACATTTGAAGATCTTCTTTGC 2400
Qy 2428 TTTGCAATCAAGTTGCCAAAGGAATGGAATTTCTGGAATTTAAAGTCGTTGTTACAGA 2487
Db 2401 TTTGCAATCAAGTTGCCAAAGGAATGGAATTTCTGGAATTTAAAGTCGTTGTTACAGA 2460
Qy 2488 GACCTGGCCGCCAGGAACCTGCTTGTCAACCCAGGAAAGTGGTGAAGATATGTGACTTT 2547
Db 2461 GACCTGGCCGCCAGGAACTGCTTGTCAACCCAGGAAAGTGGTGAAGATATGTGACTTT 2520
Qy 2548 GGATTGGCTCGAGATATCATAGATGATTCCAACTATGTTGTGACGGGGCAATGCCGCTCTG 2607
Db 2521 GGATTGGCTCGAGATATCATAGATGATTCCAACTATGTTGTGACGGGGCAATGCCGCTCTG 2580
Qy 2608 CCTGTAAATGATGGCCCCCGAAGAGCTGTTTGAAGGCATCTACACCATTAAGATGAT 2667
Db 2581 CCTGTAAATGATGGCCCCCGAAGAGCTGTTTGAAGGCATCTACACCATTAAGATGAT 2640
Qy 2668 GTCTGTCATATGGAATATTTACTGTGGGAAATCTTCTCACTTGGTGTGAATCCTTACCCT 2727
Db 2641 GTCTGTCATATGGAATATTTACTGTGGGAAATCTTCTCACTTGGTGTGAATCCTTACCCT 2700
Qy 2728 GGCATTCCGGTTGATGCTAACTTCTACAACTGATTCAAAAATGGATTTAAATGGATCAG 2787
Db 2701 GGCATTCCGGTTGATGCTAACTTCTACAACTGATTCAAAAATGGATTTAAATGGATCAG 2760
Qy 2788 CCATTTTATGCTACAGAGAAATATACATTAATGCAATCCTGCTGGGCTTTTGACTCA 2847
Db 2761 CCATTTTATGCTACAGAGAAATATACATTAATGCAATCCTGCTGGGCTTTTGACTCA 2820
Qy 2848 AGGAAACGGCCATCCTTCCCTAATTTGACTTCGTTTTTTAGGATGTCAGCTGGCAGATGCA 2907
Db 2821 AGGAAACGGCCATCCTTCCCTAATTTGACTTCGTTTTTTAGGATGTCAGCTGGCAGATGCA 2880
Qy 2908 GAAGAGCGATGATCAGAAATGCGATGGCGGTGTTTCGGAATGTCCTCACCTACCAA 2967
Db 2881 GAAGAGCGATGATCAGAAATGCGATGGCGGTGTTTCGGAATGTCCTCACCTACCAA 2940
Qy 2968 AACAGCGCA 2976
Db 2941 AACAGCGCA 2949
```

Search completed: August 27, 2003, 16:01:41
Job time : 1052.15 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 28, 2003, 01:02:19 ; Search time 1021.5 Seconds
(without alignments)
7863.095 Million cell updates/sec

Title: US-09-919-408a-3

Perfect score: 3501

Sequence: 1 CGAGGGCGGATCCGAGGCT.....AAAAAAAAAAAAAAAAAAAA 3501

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1533700 seqs, 1147125425 residues

Total number of hits satisfying chosen parameters: 3067400

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 16: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 17: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3501	100.0	3501	9	US-09-919-408-3
2	3501	100.0	3501	10	US-09-872-136-3
3	3083.2	88.1	3120	10	US-09-982-610-22
4	2227.2	63.6	3453	9	US-09-919-408-1
5	2227.2	63.6	3453	10	US-09-872-136-1
6	633	18.1	645	10	US-09-796-692-7538
7	633	18.1	645	14	US-10-040-862-7538
8	608.4	17.4	610	10	US-09-796-692-8085
9	608.4	17.4	610	14	US-10-040-862-8085
c 10	606.4	17.3	608	10	US-09-796-692-8132
c 11	606.4	17.3	608	14	US-10-040-862-8132
12	605	17.3	605	10	US-09-796-692-8927
13	605	17.3	605	14	US-10-040-862-8927
14	598.4	17.1	601	10	US-09-796-692-8273
15	598.4	17.1	601	14	US-10-040-862-8273
16	597	17.1	597	10	US-09-796-692-7976

17	597	17.1	597	14	US-10-040-862-7976	Sequence 7976, Ap
18	595.4	17.0	597	10	US-09-796-692-9010	Sequence 9010, Ap
19	595.4	17.0	597	14	US-10-040-862-9010	Sequence 9010, Ap
20	593.2	16.9	598	10	US-09-796-692-8566	Sequence 8566, Ap
21	593.2	16.9	598	14	US-10-040-862-8566	Sequence 8566, Ap
22	590.8	16.9	594	10	US-09-796-692-8135	Sequence 8135, Ap
23	590.8	16.9	594	14	US-10-040-862-8135	Sequence 8135, Ap
24	590.4	16.9	605	10	US-09-796-692-8047	Sequence 8047, Ap
25	590.4	16.9	605	14	US-10-040-862-8047	Sequence 8047, Ap
26	583.8	16.7	589	10	US-09-796-692-8630	Sequence 8630, Ap
27	583.8	16.7	589	14	US-10-040-862-8630	Sequence 8630, Ap
28	562.4	16.1	564	10	US-09-796-692-9025	Sequence 9025, Ap
29	562.4	16.1	564	14	US-10-040-862-9025	Sequence 9025, Ap
30	554.4	15.8	556	10	US-09-796-692-7883	Sequence 7883, Ap
31	554.4	15.8	556	14	US-10-040-862-7883	Sequence 7883, Ap
32	541	15.5	550	10	US-09-796-692-8494	Sequence 8494, Ap
33	541	15.5	550	14	US-10-040-862-8494	Sequence 8494, Ap
c 34	539	15.4	539	10	US-09-796-692-8255	Sequence 8255, Ap
c 35	539	15.4	539	14	US-10-040-862-8255	Sequence 8255, Ap
36	510	14.6	510	10	US-09-796-692-8268	Sequence 8268, Ap
37	510	14.6	510	14	US-10-040-862-8268	Sequence 8268, Ap
38	502.8	14.4	508	10	US-09-796-692-8032	Sequence 8032, Ap
39	502.8	14.4	508	14	US-10-040-862-8032	Sequence 8032, Ap
40	491.4	14.0	497	10	US-09-796-692-8408	Sequence 8408, Ap
41	491.4	14.0	497	14	US-10-040-862-8408	Sequence 8408, Ap
42	478.4	13.7	606	10	US-09-796-692-8114	Sequence 8114, Ap
43	478.4	13.7	606	14	US-10-040-862-8114	Sequence 8114, Ap
44	446.4	12.8	581	10	US-09-796-692-9031	Sequence 9031, Ap
45	446.4	12.8	581	14	US-10-040-862-9031	Sequence 9031, Ap

ALIGNMENTS

RESULT 1

US-09-919-408-3
; Sequence 3: Application US/09919408
; Patent No. US20020072077A1
; GENERAL INFORMATION:
; APPLICANT: Lemischka, Ihor R.
; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL RECEPTORS AND THEIR LIGANDS
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Imclone Systems Incorporated
; STREET: 180 Varick Street
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10014
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: 31-Jul-2001
; APPLICATION NUMBER: US/09/919,408
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/977,451
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US 07/906,397
; FILING DATE: 26-JUN-1992
; APPLICATION NUMBER: US PCT/US92/05401
; FILING DATE: 26-JUN-1992
; APPLICATION NUMBER: TW 81102961
; FILING DATE: 15-APR-1992
; APPLICATION NUMBER: US PCT/US92/02750
; FILING DATE: 02-APR-1992
; APPLICATION NUMBER: US 07/813,593
; FILING DATE: 24-DEC-1991
; APPLICATION NUMBER: US 07/793,065

FILING DATE: 15-NOV-1991
APPLICATION NUMBER: US 07/728,913
FILING DATE: 28-JUN-1991
APPLICATION NUMBER: US 07/679,666
FILING DATE: 02-APR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Feit, Irving N.
REGISTRATION NUMBER: 28,601
REFERENCE/DOCKET NUMBER: LEM-3-7P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-645-1405
TELEFAX: 212-645-2054
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 3501 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
FEATURE:
NAME/KEY: CDS
LOCATION: 58..3039
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 139..3036
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 58..138
SEQUENCE DESCRIPTION: SEQ ID NO: 3:

US-09-919-408-3
Query Match 100.0%; Score 3501; DB 9; Length 3501;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	CGAGGGCGCATCCGAGGCTGGGCGCGCCCTGGGGACCCCGGGCTCCGGAGCCCATG	60
Db	1	CGAGGGCGCATCCGAGGCTGGGCGCGCCCTGGGGACCCCGGGCTCCGGAGCCCATG	60
Qy	61	CCGGCGTTGGCGCGAGCGGGACCGTGGCGCTGCTGTTGTTTCTGCAATGATA	120
Db	61	CCGGCGTTGGCGCGAGCGGGACCGTGGCGCTGCTGTTGTTTCTGCAATGATA	120
Qy	121	TTTGGGACTATTACAAATCAAGATCTGCTGTGATCAAGTGTGTTTAAATCAATCAAG	180
Db	121	TTTGGGACTATTACAAATCAAGATCTGCTGTGATCAAGTGTGTTTAAATCAATCAAG	180
Qy	181	AACAATGATTTCATCAGTGGGGAAGTCATCATATCCCATGATCAGATCCCGGAA	240
Db	181	AACAATGATTTCATCAGTGGGGAAGTCATCATATCCCATGATCAGATCCCGGAA	240
Qy	241	GACCTCGGGTGGCGTTGAGACCCAGAGCTCAGGAGCAGTGTACGAGCTGCCGCTGTG	300
Db	241	GACCTCGGGTGGCGTTGAGACCCAGAGCTCAGGAGCAGTGTACGAGCTGCCGCTGTG	300
Qy	301	GAAGTGTATGATCTGCTCCATCAGTCAAGTGTGTCGATGCCCGAGGGAACATT	360
Db	301	GAAGTGTATGATCTGCTCCATCAGTCAAGTGTGTCGATGCCCGAGGGAACATT	360
Qy	361	TCCTGTCTCTGGGCTTTAAGCACAGCTCCCTGAAATGCCAGCCACATTTTGAATACAA	420
Db	361	TCCTGTCTCTGGGCTTTAAGCACAGCTCCCTGAAATGCCAGCCACATTTTGAATACAA	420
Qy	421	AACAGAGAGTGTGTTTCATGCTCATTTTGAATGACAGAAACCCAGCTGGAGATAC	480
Db	421	AACAGAGAGTGTGTTTCATGCTCATTTTGAATGACAGAAACCCAGCTGGAGATAC	480
Qy	481	CTACTTTTATTACAGTGAAGCTACCAATACCAATATTTTACAGTGAATGAAGA	540
Db	481	CTACTTTTATTACAGTGAAGCTACCAATACCAATATTTTACAGTGAATGAAGA	540

Qy	541	AATACCTCTGCTTTACACATTAAGAGACCTTACTTTAGAAAATGGAACCCAGGACGCC	600
Db	541	AATACCTCTGCTTTACACATTAAGAGACCTTACTTTAGAAAATGGAACCCAGGACGCC	600
Qy	601	CTGCTCTGCATATCTGAGAGCGTTCCAGAGCCCATCGTGAATGGGTGCTTTGCGATTCA	660
Db	601	CTGCTCTGCATATCTGAGAGCGTTCCAGAGCCCATCGTGAATGGGTGCTTTGCGATTCA	660
Qy	661	CAGGGGAAAGCTGTAAAGAAAGAGTCCAGCTGTTTAAAGAGGAGGAAAAGTGTCTT	720
Db	661	CAGGGGAAAGCTGTAAAGAAAGAGTCCAGCTGTTTAAAGAGGAGGAAAAGTGTCTT	720
Qy	721	CATGAATTTTGGGACCGGACATAAGTGTCTGCCAGAAATGAATCTGGGACGGAATGC	780
Db	721	CATGAATTTTGGGACCGGACATAAGTGTCTGCCAGAAATGAATCTGGGACGGAATGC	780
Qy	781	ACCAGGCTGTTCAATATAGATCTAAATCAAATCTCTCAGACCACATGTCACCAATATTT	840
Db	781	ACCAGGCTGTTCAATATAGATCTAAATCAAATCTCTCAGACCACATGTCACCAATATTT	840
Qy	841	CTTAAAGTAGGGGAACCTTTATGATAAGGTGCAAGCTGTTTCAATGTAACCATGGAATC	900
Db	841	CTTAAAGTAGGGGAACCTTTATGATAAGGTGCAAGCTGTTTCAATGTAACCATGGAATC	900
Qy	901	GGGCTCACCTGGGAAATTAGAAAACAAAGCACTCAGAGGGGCAACTACTTTGAGATGAGT	960
Db	901	GGGCTCACCTGGGAAATTAGAAAACAAAGCACTCAGAGGGGCAACTACTTTGAGATGAGT	960
Qy	961	ACCTATTTCAACAAACAGAACTATGATACGATGTTCTGTTTCTTTGTATCATCAGTGGA	1020
Db	961	ACCTATTTCAACAAACAGAACTATGATACGATGTTCTGTTTCTTTGTATCATCAGTGGA	1020
Qy	1021	AGAAACGACACCGGATCTACACTGTTCTCTTCAAGCATCCAGTCAATCAGCTTTG	1080
Db	1021	AGAAACGACACCGGATCTACACTGTTCTCTTCAAGCATCCAGTCAATCAGCTTTG	1080
Qy	1081	GTTACCATCTAGAAAAGGATTTAATGCTACCAATTTCAAGTGAAGATTTAAGAAAT	1140
Db	1081	GTTACCATCTAGAAAAGGATTTAATGCTACCAATTTCAAGTGAAGATTTAAGAAAT	1140
Qy	1141	GACCAATATGAAGAGTTTCTGTTCTGTCAGGTTTAAAGCTTACCCACAAATCAGATGT	1200
Db	1141	GACCAATATGAAGAGTTTCTGTTCTGTCAGGTTTAAAGCTTACCCACAAATCAGATGT	1200
Qy	1201	ACGTGGACCTTCTCTCGAAAATCATTTCTTGTGAGCAAAAGGCTCTTTGATAACGGATAC	1260
Db	1201	ACGTGGACCTTCTCTCGAAAATCATTTCTTGTGAGCAAAAGGCTCTTTGATAACGGATAC	1260
Qy	1261	AGCATATCCAAGTTTTCGAATCATAGCACAGCCAGGAGAAATATATATTCATGCGAGAA	1320
Db	1261	AGCATATCCAAGTTTTCGAATCATAGCACAGCCAGGAGAAATATATATTCATGCGAGAA	1320
Qy	1321	AATGATGATGCCAAATTTTACCAAAATGTTCCGCTGAATATAAGAGGAAACCTCAAGTG	1380
Db	1321	AATGATGATGCCAAATTTTACCAAAATGTTCCGCTGAATATAAGAGGAAACCTCAAGTG	1380
Qy	1381	CTGCGAAGCATCGGCAAGTCTGAGGCTCTCTGTTCTCGGATGATACCCATTTACCATCT	1440
Db	1381	CTGCGAAGCATCGGCAAGTCTGAGGCTCTCTGTTCTCGGATGATACCCATTTACCATCT	1440
Qy	1441	TGGAATGGAAGAGTGTTCAGACAGTCTCCCAATCTGCAAGAGAGATCACAGAGGA	1500
Db	1441	TGGAATGGAAGAGTGTTCAGACAGTCTCCCAATCTGCAAGAGAGATCACAGAGGA	1500
Qy	1501	GTCTGGAATGGAAGGCTTAACAGAAAGTGTTCGACAGTGGGTGTCGAGCAGTACTCTA	1560
Db	1501	GTCTGGAATGGAAGGCTTAACAGAAAGTGTTCGACAGTGGGTGTCGAGCAGTACTCTA	1560
Qy	1561	AACATGATGAGGAGTCTTCTGCTCAAGTGTGTCATACAAATTCCTCTGGC	1620
Db	1561	AACATGATGAGGAGTCTTCTGCTCAAGTGTGTCATACAAATTCCTCTGGC	1620

1621 ACATCTTGTCAGAGATCCCTTTTAAACTCTCCAGGCCCTTCCCTTTTCATCCAGACAAC 1680
1621 ACATCTTGTCAGAGATCCCTTTTAAACTCTCCAGGCCCTTCCCTTTTCATCCAGACAAC 1680
1681 ATCTCAATCTATGCAACAATGGTGTGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1740
1681 ATCTCAATCTATGCAACAATGGTGTGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1740
1741 ATTTGTCAACAAGTACAAAAGCAATTTAGGTATGAAAAGCCAGCTACAGATGGTACAGGTG 1800
1741 ATTTGTCAACAAGTACAAAAGCAATTTAGGTATGAAAAGCCAGCTACAGATGGTACAGGTG 1800
1801 ACCGGCTCCTCAGATATGATGATCTCTACGTGTTGATTTTCAAGAAATATGATATGATCTC 1860
1801 ACCGGCTCCTCAGATATGATGATCTCTACGTGTTGATTTTCAAGAAATATGATATGATCTC 1860
1861 AAATGGAGTTTCCAAAGAGAAAATTTAGAGTTTGGGAAGGTACTAGGATCAGGTGCTTTT 1920
1921 GGAAGAGTATGAAACGCAAGCTTATGGAATTTAGCAAAACAGAGGTCTCAATCCAGGTT 1980
1921 GGAAGAGTATGAAACGCAAGCTTATGGAATTTAGCAAAACAGAGGTCTCAATCCAGGTT 1980
1981 GCCGTCAAAATGCTGAAAGAAAAGCAGACAGCTCTGAAAGAGAGGCACCTCATGTCAGAA 2040
1981 GCCGTCAAAATGCTGAAAGAAAAGCAGACAGCTCTGAAAGAGAGGCACCTCATGTCAGAA 2040
2041 CTCAAGATGATGACCCAGCTGGGAAGCCACAGAAATTTGTGAACCTGCTGGGGGGCGTGC 2100
2041 CTCAAGATGATGACCCAGCTGGGAAGCCACAGAAATTTGTGAACCTGCTGGGGGGCGTGC 2100
2101 ACATGTCAGGACCAATTTACTTGTATTTTGAATTTGATGTTGCTATGTTGATGTTCTCAAC 2160
2101 ACATGTCAGGACCAATTTACTTGTATTTTGAATTTGATGTTGCTATGTTGATGTTCTCAAC 2160
2161 TATCTAAGAGTAAAGAGAAAATTTCAAGGACTTTGGAAGAGGATTTTCAAGGAACAC 2220
2161 TATCTAAGAGTAAAGAGAAAATTTCAAGGACTTTGGAAGAGGATTTTCAAGGAACAC 2220
2221 AATTTCAAGTTTACCCCACTTTCCAAATCATATCCAAATTTCCAGCATGCTGTTCAAGA 2280
2221 AATTTCAAGTTTACCCCACTTTCCAAATCATATCCAAATTTCCAGCATGCTGTTCAAGA 2280
2281 GAAGTTTCAAGTACACCCGAGCTCGGATCAAATCTCAGGGCTTCATGGAATTCATTTTCAAC 2340
2281 GAAGTTTCAAGTACACCCGAGCTCGGATCAAATCTCAGGGCTTCATGGAATTCATTTTCAAC 2340
2341 TCTGAAGATGAATTTGAATATGAAAACCAAAAAGGCTGGAAGAGAGGAGGACTTGAAT 2400
2341 TCTGAAGATGAATTTGAATATGAAAACCAAAAAGGCTGGAAGAGAGGAGGACTTGAAT 2400
2401 GTGCTTACATTTGAAGATCTTTGCTGTTGATTTGATTTGATTTGATTTGATTTGATTTT 2460
2401 GTGCTTACATTTGAAGATCTTTGCTGTTGATTTGATTTGATTTGATTTGATTTGATTTT 2460
2461 CTGGAATTTAAGTCTGTTTCAAGAGACTGCGCCAGGAACTGCTTGTCAACCCAC 2520
2461 CTGGAATTTAAGTCTGTTTCAAGAGACTGCGCCAGGAACTGCTTGTCAACCCAC 2520
2521 GGAAGAGTGTGAAGATATGATGATTTGATTTGATTTGATTTGATTTGATTTGATTTTCAAC 2580
2521 GGAAGAGTGTGAAGATATGATGATTTGATTTGATTTGATTTGATTTGATTTGATTTTCAAC 2580
2581 TATGTTGTCAGGGGCAATCCGCTGCTGCTGTAATGGAATGGATGGCCCGGAAAGCCTGTTT 2640
2581 TATGTTGTCAGGGGCAATCCGCTGCTGCTGTAATGGAATGGATGGCCCGGAAAGCCTGTTT 2640
2641 GAAGGATCTACACCATTAAGAGTATGATGTTGATTTGATTTGATTTGATTTGATTTTCAAC 2700
2641 GAAGGATCTACACCATTAAGAGTATGATGTTGATTTGATTTGATTTGATTTGATTTTCAAC 2700
2701 TTCTCACTGGTGTGAATCCTTACCTGGCATTTCCGCTGATGCTAACTTCTACAAACTG 2760

2701 TTCTCACTTGGTGTGAATCCTTACCTGGCATTTCCGTTGATGCTAATCTTCTACAACTG 2760
2761 ATTCAAAATGGATTTAAATGGATCAGCCATTTTATGCTACAGAAAGAAATATACATTATA 2820
2761 ATTCAAAATGGATTTAAATGGATCAGCCATTTTATGCTACAGAAAGAAATATACATTATA 2820
2821 ATGCAATCCTCTGGGCTTTTGAATCAAGAAAGCCGATCTTCCCTTAATTTGACTTCG 2880
2821 ATGCAATCCTCTGGGCTTTTGAATCAAGAAAGCCGATCTTCCCTTAATTTGACTTCG 2880
2881 TTTTGTAGGATGTCAAGTGGCAGATGCAAGAAAGCCGATGATCAGAAATGTGGATGGCCGT 2940
2881 TTTTGTAGGATGTCAAGTGGCAGATGCAAGAAAGCCGATGATCAGAAATGTGGATGGCCGT 2940
2941 GTTTCGAATGTCTCTACACCTACCAAAACAGGCGACCTTTTACGACAGAGATGGATTTG 3000
2941 GTTTCGAATGTCTCTACACCTACCAAAACAGGCGACCTTTTACGACAGAGATGGATTTG 3000
3001 GGCTACTCTCTCGGAGGCTCAGGTCAAGATTTCTAGAGGAACAAATTTAGTTTAAAG 3060
3001 GGCTACTCTCTCGGAGGCTCAGGTCAAGATTTCTAGAGGAACAAATTTAGTTTAAAG 3060
3061 ACTTCATCCTCCACCTATCCCTAAAGGCTGTAGATTACCAAAACAGATTAATTTTCAAT 3120
3061 ACTTCATCCTCCACCTATCCCTAAAGGCTGTAGATTACCAAAACAGATTAATTTTCAAT 3120
3121 CACTAAAAGAAAATCTATTATCACTGCTGCTTCCACAGACTTTTCTCTAGAAGCCGTCT 3180
3121 CACTAAAAGAAAATCTATTATCACTGCTGCTTCCACAGACTTTTCTCTAGAAGCCGTCT 3180
3181 GCGTTACTCTGTTTCAAGGGAATTTTGTAAAATCAATCATCTCTGTCACAAGGCGAG 3240
3181 GCGTTACTCTGTTTCAAGGGAATTTTGTAAAATCAATCATCTCTGTCACAAGGCGAG 3240
3241 GAGGCTGATATGAACTTTTATGGAGCATGATGTCATCCAGGCTTCTCAGGCGG 3300
3241 GAGGCTGATATGAACTTTTATGGAGCATGATGTCATCCAGGCTTCTCAGGCGG 3300
3301 GCTTGAGTGAATTTGCTGACCTGAAGTACAGTATATTCTTGTAAAATACATAAAACAAAGC 3360
3301 GCTTGAGTGAATTTGCTGACCTGAAGTACAGTATATTCTTGTAAAATACATAAAACAAAGC 3360
3361 ATTTGCTAAGGAGAGCTAATATGATTTTAAAGTCTATGTTTAAATTAATATGTAAA 3420
3361 ATTTGCTAAGGAGAGCTAATATGATTTTAAAGTCTATGTTTAAATTAATATGTAAA 3420
3421 TTTTTCAGCTATTAGTATATATTTTATGGTGGGAATTAATTTCTCTACAGAAAAA 3480
3421 TTTTTCAGCTATTAGTATATATTTTATGGTGGGAATTAATTTCTCTACAGAAAAA 3480
3481 AAAAAAAAAAAAAAAAAAAAAA 3501
3481 AAAAAAAAAAAAAAAAAAAAAA 3501

RESULT 2

US-09-872-136-3

; Sequence 3, Application US/09872136

; Patent No. US20020119545A1

; GENERAL INFORMATION:

; APPLICANT: Lemischka, Ihor R.

; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL

; RECEPTORS AND THEIR LIGANDS

; NUMBER OF SEQUENCES: 10

; CORRESPONDENCE ADDRESS:

; ADDRESS: InClone Systems Incorporated

; STREET: 180 Varick Street

; CITY: New York

; STATE: New York

; COUNTRY: U.S.A.

; ZIP: 10014

; COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/872,136
FILING DATE: 01-Jun-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/208,786
FILING DATE: <Unknown>
APPLICATION NUMBER: US/09/021,324
FILING DATE: <Unknown>
APPLICATION NUMBER: US/07/977,451
FILING DATE: 1992-11-19
APPLICATION NUMBER: US 07/906,397
FILING DATE: 26-JUN-1992
APPLICATION NUMBER: US PCT/US92/05401
FILING DATE: 26-JUN-1992
APPLICATION NUMBER: TW 81102961
FILING DATE: 15-APR-1992
APPLICATION NUMBER: US PCT/US92/02750
FILING DATE: 02-APR-1992
APPLICATION NUMBER: US 07/813,593
FILING DATE: 24-DEC-1991
APPLICATION NUMBER: US 07/793,065
FILING DATE: 15-NOV-1991
APPLICATION NUMBER: US 07/728,913
FILING DATE: 28-JUN-1991
APPLICATION NUMBER: US 07/679,666
FILING DATE: 02-APR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Feit, Irving N.
REGISTRATION NUMBER: 28,601
REFERENCE/DOCKET NUMBER: LEM-3-7P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-645-1405
TELEFAX: 212-645-2054
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 3501 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
FEATURE:
NAME/KEY: CDS
LOCATION: 58..3039
FEATURE:
NAME/KEY: mat.peptide
LOCATION: 139..3036
FEATURE:
NAME/KEY: sig.peptide
LOCATION: 58..138
SEQUENCE DESCRIPTION: SEQ ID NO: 3:

121	TTTGGGACTATTACAAATCAAGATCTGCTGTGATCAAGTGTGTTTAAATCAATCAAG	180
181	AACAATGATTTCATCAGTGGGAAAGTTCATCATATATCCCATGATGATCAGAAATCCCGGAA	240
181	AACAATGATTTCATCAGTGGGAAAGTTCATCATATATCCCATGATGATCAGAAATCCCGGAA	240
241	GACCTGGGTGTGGTGTGAGACCCAGAGCTCAGGAGCAGTGTACGAAGCTGCCGCTGTG	300
241	GACCTGGGTGTGGTGTGAGACCCAGAGCTCAGGAGCAGTGTACGAAGCTGCCGCTGTG	300
301	GAGTGGATGATCTGCTTCCATCACACTGCAAGTGTGCTGCAATGCTCCAGAGGAAACATT	360
301	GAGTGGATGATCTGCTTCCATCACACTGCAAGTGTGCTGCAATGCTCCAGAGGAAACATT	360
361	TCCTGTCTCTGGGTCTTTTAAAGCACAGCTCCCTGAAATTCAGCCACATTTTGAATTTACAA	420
361	TCCTGTCTCTGGGTCTTTTAAAGCACAGCTCCCTGAAATTCAGCCACATTTTGAATTTACAA	420
421	AACAGAGGAGTTGTTTCCATGGTCAATTTGAAATGACAGAAACCCAGCTGGGAGATAC	480
421	AACAGAGGAGTTGTTTCCATGGTCAATTTGAAATGACAGAAACCCAGCTGGGAGATAC	480
481	CTACTTTTATTCAGAGTGAAGTACCAATTTACACAATATTTTACAGTCAAGTATAGA	540
481	CTACTTTTATTCAGAGTGAAGTACCAATTTACACAATATTTTACAGTCAAGTATAGA	540
541	AATACCTCTGCTTTACACATTTAAGAGACCTTACTTTTAAAGAAATGGAAGAACCCAGGACGC	600
541	AATACCTCTGCTTTACACATTTAAGAGACCTTACTTTTAAAGAAATGGAAGAACCCAGGACGC	600
601	CTGCTCTGCATATCTGAGAGCGTTCCAGAGCCGATCGTGAATGGGTGCTTTGCAATTCA	660
601	CTGCTCTGCATATCTGAGAGCGTTCCAGAGCCGATCGTGAATGGGTGCTTTGCAATTCA	660
661	CAGGGGAAAGCTGTAAAGAGAAAGTCCAGCTGCTTCTTAAAGAGGAGGAAAGTGCCT	720
661	CAGGGGAAAGCTGTAAAGAGAAAGTCCAGCTGCTTCTTAAAGAGGAGGAAAGTGCCT	720
721	CATGAATTTTGGGACGGACATAAGGTGCTGCCAGAAATGAATCTGGGAGGAAATGC	780
721	CATGAATTTTGGGACGGACATAAGGTGCTGCCAGAAATGAATCTGGGAGGAAATGC	780
781	ACCAGGCTGTTCAATAGATCTAAATCAAACTCCTCAGACCAATGTCCTCAATTTT	840
781	ACCAGGCTGTTCAATAGATCTAAATCAAACTCCTCAGACCAATGTCCTCAATTTT	840
841	CTTAAAGTAGGGAAACCCCTTATGGATAAGGTGCAAGCTGTTCAATGGAATTC	900
841	CTTAAAGTAGGGAAACCCCTTATGGATAAGGTGCAAGCTGTTCAATGGAATTC	900
901	GGGCTCACCTGGGAAATTTAGAAAACAAAGACCTCCAGAGGGGCAACTACTTTGAGATGAGT	960
901	GGGCTCACCTGGGAAATTTAGAAAACAAAGACCTCCAGAGGGGCAACTACTTTGAGATGAGT	960
961	ACCTATTCAACAAACAGAACTATGATACGGATCTGTTTGTCTTTTGTATCATCAGTGGA	1020
961	ACCTATTCAACAAACAGAACTATGATACGGATCTGTTTGTCTTTTGTATCATCAGTGGA	1020
1021	AGAAACGACACCGGATACCTACCTTGTCTCTTCAAGGATCCAGTCAATCAGCTTTG	1080
1021	AGAAACGACACCGGATACCTACCTTGTCTCTTCAAGGATCCAGTCAATCAGCTTTG	1080
1081	GTTACCATCTGAGAAAGGATTTTAAATGCTCACTAAATCAAGTGAAGATTTAGAAAT	1140
1081	GTTACCATCTGAGAAAGGATTTTAAATGCTCACTAAATCAAGTGAAGATTTAGAAAT	1140
1141	GACCAATATGAAGAGTTTGTGTTTCTGCTAGGTTTAAAGCTACCCACAAATCAGATGT	1200
1141	GACCAATATGAAGAGTTTGTGTTTCTGCTAGGTTTAAAGCTACCCACAAATCAGATGT	1200
1201	ACGTGGACCTTCTCTCGAAATCAATTTCTGCTGAGCAAAAGGGTCTTTGATAACGATAC	1260

Query Match 100.0%; Score 3501; DB 10; Length 3501;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGAGGCGCATCCGAGGCTGGCGCGCCCTCGGGGACCCCGGGCTCCGGAGGCCATG 60
DB 1 CGAGGCGCATCCGAGGCTGGCGCGCCCTCGGGGACCCCGGGCTCCGGAGGCCATG 60
QY 61 CCGGCGTTGGCGCGGACCGGACCGTGGCTGCTGCTGTTTCTGCAATGATA 120
DB 61 CCGGCGTTGGCGCGGACCGGACCGTGGCTGCTGCTGTTTCTGCAATGATA 120
QY 121 -TTTGGGACTATTACAAATCAAGATCTGCTGTGATCAAGTGTGTTTAAATCAATCAAG 180

Db	1201	ACGTGACCTTCTCGAAATCATTTCTTGTGAGCAAAAGGCTCTTGATAACGGATAC	1260
Qy	1261	AGCATATCCAAAGTTTGGCAATCATATAGCACACAGCCAGGAGAAATATATATTCATTCAGAA	1320
Db	1261	AGCATATCCAAAGTTTGGCAATCATATAGCACACAGCCAGGAGAAATATATATTCATTCAGAA	1320
Qy	1321	AATGATGATGCCCAATTTACCAAAATGTTCCAGCTGAATATAAGAGAAACCTCAAGTG	1380
Db	1321	AATGATGATGCCCAATTTACCAAAATGTTCCAGCTGAATATAAGAGAAACCTCAAGTG	1380
Qy	1381	CTCGAGAAGCATCGGCAAGTCAGGCTCTCTGTTCTCGGATGGATACCCATTACCATCT	1440
Db	1381	CTCGAGAAGCATCGGCAAGTCAGGCTCTCTGTTCTCGGATGGATACCCATTACCATCT	1440
Qy	1441	TGACCTGGAAGAGTGTTCAGACAACTCTCCAACTGCAAGAGATCACAGAAGGA	1500
Db	1441	TGACCTGGAAGAGTGTTCAGACAACTCTCCAACTGCAAGAGATCACAGAAGGA	1500
Qy	1501	GTCTGGAATAGAAAGGCTAACAGAAAGTGTGTTGGACAGTGGGTGTCGAGCAGTACTCTA	1560
Db	1501	GTCTGGAATAGAAAGGCTAACAGAAAGTGTGTTGGACAGTGGGTGTCGAGCAGTACTCTA	1560
Qy	1561	AACATGAGTGAAGCCATAAAGGTTCTCTGCTCAAGTGTCTGTCATACAAATTCCTTTGGC	1620
Db	1561	AACATGAGTGAAGCCATAAAGGTTCTCTGCTCAAGTGTCTGTCATACAAATTCCTTTGGC	1620
Qy	1621	ACATCTGTGAGCAGTCTTTTAACTCTCGAGGCCCTTCCCTTTTCATCCAGAACAC	1680
Db	1621	ACATCTGTGAGCAGTCTTTTAACTCTCGAGGCCCTTCCCTTTTCATCCAGAACAC	1680
Qy	1681	ATCTCATCTATGCAACAAATGGTGTGTTGCTCTCTTCAATGTCGTTTAAACCTGCTA	1740
Db	1681	ATCTCATCTATGCAACAAATGGTGTGTTGCTCTCTTCAATGTCGTTTAAACCTGCTA	1740
Qy	1741	ATTGTGTCAGAAATGAAAGCAATTTAGTGTATGAAAGCCAGCTACAGATGGTACAGTG	1800
Db	1741	ATTGTGTCAGAAATGAAAGCAATTTAGTGTATGAAAGCCAGCTACAGATGGTACAGTG	1800
Qy	1801	ACCGGCTCTCAGATAATAGTACTTCTACGTTGATTTACAGAAATGAATATGATCTC	1860
Db	1801	ACCGGCTCTCAGATAATAGTACTTCTACGTTGATTTACAGAAATGAATATGATCTC	1860
Qy	1861	AAATGGGAGTTTCCAGAGAAAATTTAGAGTTTGGGAAGGTACTAGGATCAGGTCTTTT	1920
Db	1861	AAATGGGAGTTTCCAGAGAAAATTTAGAGTTTGGGAAGGTACTAGGATCAGGTCTTTT	1920
Qy	1921	GGAAAGTGTATGACGCAACAGCTTTATGGAATTAGCAAAACAGGAGTCTCAATCCAGGTT	1980
Db	1921	GGAAAGTGTATGACGCAACAGCTTTATGGAATTAGCAAAACAGGAGTCTCAATCCAGGTT	1980
Qy	1981	GCGGTCAAAATGCTGAAAGAAAAGCAGACAGCTCTGAAAAGAGGCACTCATGTAGAA	2040
Db	1981	GCGGTCAAAATGCTGAAAGAAAAGCAGACAGCTCTGAAAAGAGGCACTCATGTAGAA	2040
Qy	2041	CTCAAGTATGATGAGTGGGAAGCCAGAGAAATTTGTGAACCTGCTGGGGGGCGTGC	2100
Db	2041	CTCAAGTATGATGAGTGGGAAGCCAGAGAAATTTGTGAACCTGCTGGGGGGCGTGC	2100
Qy	2101	ACACTGTGAGGACCAATTTACTGATTTTGAATGATGTTGCTATGTTGATCTTCTCAAC	2160
Db	2101	ACACTGTGAGGACCAATTTACTGATTTTGAATGATGTTGCTATGTTGATCTTCTCAAC	2160
Qy	2161	TATCTAAGAAATGAAAGAAAATTTTACAGGACTTGGACAGAGATTTTCAAGGAACAC	2220
Db	2161	TATCTAAGAAATGAAAGAAAATTTTACAGGACTTGGACAGAGATTTTCAAGGAACAC	2220
Qy	2221	AATTCAGTTTATACCCACTTTTCCAAATCATCCAAATTCAGCATGCTGGTTCAAGA	2280
Db	2221	AATTCAGTTTATACCCACTTTTCCAAATCATCCAAATTCAGCATGCTGGTTCAAGA	2280
Qy	2281	GAAGTTTCAGATACACCCGGACTCGATCAATCTCAGGCTTCATGGGAATTCATTTCAC	2340
Db	2281	GAAGTTTCAGATACACCCGGACTCGATCAATCTCAGGCTTCATGGGAATTCATTTCAC	2340

Qy	2341	TCTGAAGATGAAATTCGAATATGAAACCAAAAAAGGCTCGAAAGAGGAGGACTTGAAT	2400
Db	2341	TCTGAAGATGAAATTCGAATATGAAACCAAAAAAGGCTCGAAAGAGGAGGACTTGAAT	2400
Qy	2401	GTGCTTACATTTGAAGATCTTCTTGTGTCATATCAAGTTGCAAAAGGAATGGAATTT	2460
Db	2401	GTGCTTACATTTGAAGATCTTCTTGTGTCATATCAAGTTGCAAAAGGAATGGAATTT	2460
Qy	2461	CTGGAATTTAAGTCGTGTGTTTCAGAGACCTGGCCGCGCAGGAACGTCCTGTGACCCAC	2520
Db	2461	CTGGAATTTAAGTCGTGTGTTTCAGAGACCTGGCCGCGCAGGAACGTCCTGTGACCCAC	2520
Qy	2521	GGGAAAGTGTGAAGATATGTCGACTTTTGGATTTGGCTCGAGATATCATGAGTATTCAC	2580
Db	2521	GGGAAAGTGTGAAGATATGTCGACTTTTGGATTTGGCTCGAGATATCATGAGTATTCAC	2580
Qy	2581	TATGTTGTTCAGGGGCAATGCCCTGTGCTGCTGTAATAATGAGTGGCCCGGAAAGCCTGTTT	2640
Db	2581	TATGTTGTTCAGGGGCAATGCCCTGTGCTGCTGTAATAATGAGTGGCCCGGAAAGCCTGTTT	2640
Qy	2641	GAAGGCATCTACACCAATTAAGAGTGTCTGCTCATATGGAATATTTACTGTGGGAATC	2700
Db	2641	GAAGGCATCTACACCAATTAAGAGTGTCTGCTCATATGGAATATTTACTGTGGGAATC	2700
Qy	2701	TTCTCACTTGGTGTGATCCTTACCTGCGCATTCGGGTGATGCTTAACAACTG	2760
Db	2701	TTCTCACTTGGTGTGATCCTTACCTGCGCATTCGGGTGATGCTTAACAACTG	2760
Qy	2761	ATTCAAAATGGAATTTAAATGGAATCAGCCATTTTATGCTACAGAAAGAAATATACATATA	2820
Db	2761	ATTCAAAATGGAATTTAAATGGAATCAGCCATTTTATGCTACAGAAAGAAATATACATATA	2820
Qy	2821	ATGCAATCTGCTGGGCTTTTGACTCAAGAAACGGCCATCTTCCCTAAATTTGACCTCG	2880
Db	2821	ATGCAATCTGCTGGGCTTTTGACTCAAGAAACGGCCATCTTCCCTAAATTTGACCTCG	2880
Qy	2881	TTTTTAGGATGTGAGTGGCAGATGCAAGAAAGCGATGTATCAGATGTGGATGGCGGT	2940
Db	2881	TTTTTAGGATGTGAGTGGCAGATGCAAGAAAGCGATGTATCAGATGTGGATGGCGGT	2940
Qy	2941	GTTCGGAATGTCTCACCTACCAAAACAGGCGACTTTTACAGAGATGTGGATGGCGGT	3000
Db	2941	GTTCGGAATGTCTCACCTACCAAAACAGGCGACTTTTACAGAGATGTGGATGGCGGT	3000
Qy	3001	GGGCTACTCTCTCGCAGGCTCAGGTGCAAGATTTCTAGAGGAAACAATTTAGTTTAAAG	3060
Db	3001	GGGCTACTCTCTCGCAGGCTCAGGTGCAAGATTTCTAGAGGAAACAATTTAGTTTAAAG	3060
Qy	3061	ACTTCATCTCTCCACCTATCCCTTAAAGGCTGTAGATTACCAAAACAGAGATTAATTTCA	3120
Db	3061	ACTTCATCTCTCCACCTATCCCTTAAAGGCTGTAGATTACCAAAACAGAGATTAATTTCA	3120
Qy	3121	CATTAAGAAATCTATTAATCACTGCTGCTTCCAGAGACTTTTCTTAGAGCCGCTCT	3180
Db	3121	CATTAAGAAATCTATTAATCACTGCTGCTTCCAGAGACTTTTCTTAGAGCCGCTCT	3180
Qy	3181	GGGTTACTCTCTGTTTCAAGGGGACTTTTGAATAATCAATCATCTGTCACAGGCGAG	3240
Db	3181	GGGTTACTCTCTGTTTCAAGGGGACTTTTGAATAATCAATCATCTGTCACAGGCGAG	3240
Qy	3241	GAGGAGCTGATAATGAACTTTTATTTGGAGCATTTGATGTCATCCAGGCGCTTCTCAGGCG	3300
Db	3241	GAGGAGCTGATAATGAACTTTTATTTGGAGCATTTGATGTCATCCAGGCGCTTCTCAGGCG	3300
Qy	3301	GCTTGAAGTGTGTGTCACCTGAGTACAGTATATTTCTTGAATATACATAAAACAAAGC	3360
Db	3301	GCTTGAAGTGTGTGTCACCTGAGTACAGTATATTTCTTGAATATACATAAAACAAAGC	3360
Qy	3361	ATTTTGTAGGAGAGCAATTAATGATTTTTTAAAGTCTATGTTTTTAAATTAATATGTA	3420
Db	3361	ATTTTGTAGGAGAGCAATTAATGATTTTTTAAAGTCTATGTTTTTAAATTAATATGTA	3420

Qy	3421	TTTTT	CAGCTATTTAGTGGATATATTTTATGGTGGGAATAAAAAATTTCTACTACGAAAAA	3480
Db	3421	TTTTT	CAGCTATTTAGTGGATATATTTTATGGTGGGAATAAAAAATTTCTACTACGAAAAA	3480
Qy	3481	AAAAAAAAAAAAAAAAAAAAA	3501	
Db	3481	AAAAAAAAAAAAAAAAAAAAA	3501	

RESULT 3
US-09-982-610-22
; Sequence 22, Application US/09982610
; Patent No. US20020146420A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; Bennett, Brian D.
; Goeddel, David
; Lee, James M.
; Matthews, William
; Tsai, Siao Ping
; Wood, William I.
; TITLE OF INVENTION: PROTEIN
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:

ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080

COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinPatin (Genentech)

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/982,610
FILING DATE: 17-Oct-2001
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/446,648
FILING DATE: 1996-MAY-23
APPLICATION NUMBER: 08/222616
FILING DATE: 04-APR-1994

ATTORNEY/AGENT INFORMATION:
NAME: Lee, Wendy M.
REGISTRATION NUMBER: 40,378
REFERENCE/DOCKET NUMBER: P0821P3PCT

TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-1994
TELEFAX: 415/952-9861
TELEX: 910/371-7168

INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 3120 base pairs
TYPE: Nucleic Acid
STRANDEDNESS: Single
TOPOLOGY: Linear

SEQUENCE DESCRIPTION: SEQ ID NO: 22:
US-09-982-610-22

	Query Match	88.1%	Score 3083.2;	DB 10;	Length 3120;
	Best Local Similarity	99.3%	Pred. No. 0;		
	Matches 3097;	Conservative 0;	Mismatches 23;	Indels 0;	Gaps 0;
Qy	58	ATGCCGCGTGTGGCGCGGACGCGGGACCGTGCGCTGCTCGTTGTTTTTTCTGCAATG	117		
Db	1	ATGAGAGCGTGTGGCGCGACGCGCGGCAGCTGCGCTGCTGTTTTTTCTGCAATG	60		
Qy	118	ATATTTGGGACTATTACAAATCAAGATCTGCCTCTGATCAAGTGTGTTTTTAATCAATCAT	177		
Db	61	ATATTTGGGACTATTACAAATCAAGATCTGCCTCTGATCAAGTGTGTTTTTAATCAATCAT	120		

Qy	178	AAGAA	CAATGATTCATCAGTGGGGAAGTGCATCATCATATATCCCATGTGTATCAGAAATCCCG	237
Db	121	AAGAA	CAATGATTCATCAGTGGGGAAGTGCATCATCATATATCCCATGTGTATCAGAAATCCCG	180
Qy	238	GAAGAC	CTCGGTGTCGGTTGAGACCCACAGAGCTCAGGAGCAGTGTACGAAGCTGCGCT	297
Db	181	GAAGAC	CTCGGTGTCGGTTGAGACCCACAGAGCTCAGGAGCAGTGTACGAAGCTGCGCT	240
Qy	298	GTGGAAG	TGGATGTATCTGCTTCATCACACTGCAAGTCTGTGTGATGCCAGGGAAC	357
Db	241	GTGGAAG	TGGATGTATCTGCTTCATCACACTGCAAGTCTGTGTGATGCCAGGGAAC	300
Qy	358	ATTTC	CTGCTCTGCGGTCTTTAAGACAGCTCCCTGAAATGCCAGGCCACATTTTGATTTA	417
Db	301	ATTTC	CTGCTCTGCGGTCTTTAAGACAGCTCCCTGAAATGCCAGGCCACATTTTGATTTA	360
Qy	418	CAAAA	CAGAGGAGTCTTTCCATGGTCAATTTTGAAATGACAGAAACCAAGCTGAGAA	477
Db	361	CAAAA	CAGAGGAGTCTTTCCATGGTCAATTTTGAAATGACAGAAACCAAGCTGAGAA	420
Qy	478	TACTCT	ACTTTTATTCAGAGTGAAGCTACCAATTAACAATATTTGTTTACAGTGAAGTATA	537
Db	421	TACTCT	ACTTTTATTCAGAGTGAAGCTACCAATTAACAATATTTGTTTACAGTGAAGTATA	480
Qy	538	AGAAA	TACCTGCTTTACATTAAGAAGACCTTACTTTAGAAAAATGGAACACAGGAC	597
Db	481	AGAAA	TACCTGCTTTACATTAAGAAGACCTTACTTTAGAAAAATGGAACACAGGAC	540
Qy	598	GCCT	TGCTGTCATATCTGAGAGCGTTCCAGAGCCGATCGTGGAAATGGGTCTTCGGAT	657
Db	541	GCCT	TGCTGTCATATCTGAGAGCGTTCCAGAGCCGATCGTGGAAATGGGTCTTCGGAT	600
Qy	658	TCACAG	GGGGAAGCTGTAAAGAAAGATCCAGCTGTTGTTAAAAAGGAGGAAAAAGTG	717
Db	601	TCACAG	GGGGAAGCTGTAAAGAAAGATCCAGCTGTTGTTAAAAAGGAGGAAAAAGTG	660
Qy	718	CTTCAT	GAAATATTTGGGACGACATAAAGTGTCTGTGCCAGAAATGAACTGGGACGGAA	777
Db	661	CTTCAT	GAAATATTTGGGACGACATAAAGTGTCTGTGCCAGAAATGAACTGGGACGGAA	720
Qy	778	TGCAC	CAGGCTGTTCACATAGATCTAAATCAAACTCCTCAGACACATTTGCCACAATTA	837
Db	721	TGCAC	CAGGCTGTTCACATAGATCTAAATCAAACTCCTCAGACACATTTGCCACAATTA	780
Qy	838	TTTCTT	AAAGTAGGGAAACCTTTATGGATAAGGTGCAAAAGCTGTTCATGTGAACCATGGA	897
Db	781	TTTCTT	AAAGTAGGGAAACCTTTATGGATAAGGTGCAAAAGCTGTTCATGTGAACCATGGA	840
Qy	898	TTCCGG	CTCACCTGGAAATTAGAAAACAAAGCACTCGAGGAGGGCAACTACTTTGAGATG	957
Db	841	TTCCGG	CTCACCTGGAAATTAGAAAACAAAGCACTCGAGGAGGGCAACTACTTTGAGATG	900
Qy	958	AGTAC	CTATTCAACAAACAGAACTATGATACGGATTCGTTTGTCTTTGTATCATCAGTG	1017
Db	901	AGTAC	CTATTCAACAAACAGAACTATGATACGGATTCGTTTGTCTTTGTATCATCAGTG	960
Qy	1018	GCAAG	AAAAACGACACCGGATACACTGTGTCCTCTTTCAAAGCATCCCAAGTCAATCAGCT	1077
Db	961	GCAAG	AAAAACGACACCGGATACACTGTGTCCTCTTTCAAAGCATCCCAAGTCAATCAGCT	1020
Qy	1078	TTGGT	TACCATCGTAGGAAGGGATTTATAAATGCTTACCAATTCAGAGTGAAGATTTATGAA	1137
Db	1021	TTGGT	TACCATCGTAGGAAGGGATTTATAAATGCTTACCAATTCAGAGTGAAGATTTATGAA	1080
Qy	1138	ATTGA	CAATATGAAGAGTTTGTGTCAGGTTTAAAGCCTTACCACAAATCAGA	1197
Db	1081	ATTGA	CAATATGAAGAGTTTGTGTCAGGTTTAAAGCCTTACCACAAATCAGA	1140
Qy	1198	TGTAC	TGGACCTTCTCTCGMAAATCATTTCTTGTGACCAAAAGGCTCTTGATACGGA	1257
Db	1141	TGTAC	TGGACCTTCTCTCGMAAATCATTTCTTGTGACCAAAAGGCTCTTGATACGGA	1200
Qy	1258	TACAG	CATATCCAAAGTTTGGCAATCAAGACACCGCAGGAGAAATATATATTCCATGCA	1317

1201	Db	 TACAGCATATCCAAAGTTTTCGAATCATATAGCACAGCCAGGAGAAATATATATTTCCATGCA	1260
1318	Qy	GAATAATGATGATGCCCAATTTTACAAAAATGTTTACCGCTGTAATATAAGAGAGAAACCTCAA	1377
1261	Db	GAATAATGATGATGCCCAATTTTACCAAAATGTTTCAOCTGTATATATAGAGAGAAACCTCAA	1320
1378	Qy	GTGCTCGCAGAAGATCGGCAAGTCTCAGGCGTCTGTGTTTCTCGGATGGATACCATTTACCA	1437
1321	Db	GTCTCGCAGAAGCTTCGGCAAGTCTCAGGCGTCTGTGTTTCTCGGATGGATACCATTTACCA	1380
1438	Qy	TCATTGGACCTGGGAAGAGTCTTCAGACAAGTCTCCCAACTGCACAGAAGAGATCACAGAA	1497
1381	Db	TCATTGGACCTGGGAAGAGTCTTCAGACAAGTCTCCCAACTGCACAGAAGAGATCACAGAA	1440
1498	Qy	GGAGTCTCGGAATAGAAAGCGCTTAACAGAAAAAGTGTGTTGGACAGTGGGTGTGCGAGCAGTACT	1557
1441	Db	GGAGTCTCGGAATAGAAAGCGCTTAACAGAAAAAGTGTGTTGGACAGTGGGTGTGCGAGCAGTACT	1500
1558	Qy	CTAAACATAGTGAAGGCCATAAAAGGGTTCCTGGTCAAGTGTGTGTGCATACAAATTCCTCTT	1617
1501	Db	CTAAACATAGTGAAGGCCATAAAAGGGTTCCTGGTCAAGTGTGTGTGCATACAAATTCCTCTT	1560
1618	Qy	GGCACATCTTTGTGAGACGATCCTTTTAAACCTCCAGGCCCTTCCCTTTCATCCCAAGAC	1677
1561	Db	GGCACATCTTTGTGAGACGATCCTTTTAAACCTCCAGGCCCTTCCCTTTCATCCCAAGAC	1620
1678	Qy	AACATCTCATTTCTATGCAACAAATTTGGTGTGTTGTCTCTCTTCTATGTGCTGTTTAAACCTTG	1737
1621	Db	AACATCTCATTTCTATGCAACAAATTTGGTGTGTTGTCTCTCTCTTCTATGTGCTGTTTAAACCTTG	1680
1738	Qy	CTAATTTGTGCACAAAGTACAAAAAGCAATTTAGGTATGAAAGCCAGCTACAGATGGTACAG	1797
1681	Db	CTAATTTGTGCACAAAGTACAAAAAGCAATTTAGGTATGAAAGCCAGCTACAGATGGTACAG	1740
1798	Qy	GTGACCGGCTCTCAGATAATAGAGTACTTCTAGTTGATTTTCAGAGAAATATGAATATGAT	1857
1741	Db	GTGACCGGATCCTCAGATTATGAGTACTTCTAGTTGATTTTCAGAGAAATATGAATATGAT	1800
1858	Qy	CTCAAAATGGGAGTTTCCACAGAGAAAATTTTAGAGTTTGGGAAGGTACTTAGGATCAGGTGCT	1917
1801	Db	CTCAAAATGGGAGTTTCCACAGAGAAAATTTTAGAGTTTGGGAAGGTACTTAGGATCAGGTGCT	1860
1918	Qy	TTTGGAAAAGTGATGAACGCAACAGCTTATGGAATTTAGCAAAAACAGAGGTCTCAATCCAG	1977
1861	Db	TTTGGAAAAGTGATGAACGCAACAGCTTATGGAATTTAGCAAAAACAGAGGTCTCAATCCAG	1920
1978	Qy	GTGTCGGTCAAAATGCTGAAAGAAAAGCAGACAGCTCTGAAAAGAGGGCACTCATGTCA	2037
1921	Db	GTATACCGTCAAAATGCTGAAAGAAAAGCAGACAGCTCTGAAAAGAGGGCACTCATGTCA	1980
2038	Qy	GAACTCAAGATGATGACCCAGCTGGGAAGCCACGAGAAATTTCTGAACTGCTGGGGGGG	2097
1981	Db	GAACTCAAGATGATGACCCAGCTGGGAAGCCACGAGAAATTTCTGAACTGCTGGGGGGG	2040
2098	Qy	TGCACACTGTCCAGGACCAATTTACTTGATTTTTGAATCTGTGTTCTATGTTGATCTTCTC	2157
2041	Db	TGCACACTGTCCAGGACCAATTTACTTGATTTTTGAATCTGTGTTCTATGTTGATCTTCTC	2100
2158	Qy	AATATCTAAGAGATTAAGAGAAAAATTTTCAAGGACTTTGGAAGAGAGATTTTCAAGGAA	2217
2101	Db	AATATCTAAGAGATTAAGAGAAAAATTTTCAAGGACTTTGGAAGAGAGATTTTCAAGGAA	2160
2218	Qy	CACAAATTCAGTTTTTACCCCACTTTCCATCATCCTCAAAATTTCCAGCATGCCCTGTTCA	2277
2161	Db	CACAAATTCAGTTTTTACCCCACTTTCCATCATCCTCAAAATTTCCAGCATGCCCTGTTCA	2220
2278	Qy	AGAGAAGTTCAGATACACCCGGACTCGGATCAAAATCTCAGGGCTTTCATGGGAATTCATTT	2337
2221	Db	AGAGAAGTTCAGATACACCCGGACTCGGATCAAAATCTCAGGGCTTTCATGGGAATTCATTT	2280
2338	Qy	CACCTCTGAAGATGAAATTTGAATATGAAAAACCAAAAGGCTTGGGAAGAGAGGAGGACTTG	2397

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/919,408

FILING DATE: 31-Jul-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/977,451

FILING DATE: <Unknown>

APPLICATION NUMBER: US 07/906,397

FILING DATE: 26-JUN-1992

APPLICATION NUMBER: US PCT/US92/05401

FILING DATE: 26-JUN-1992

APPLICATION NUMBER: TW 81102961

FILING DATE: 15-APR-1992

APPLICATION NUMBER: US PCT/US92/02750

FILING DATE: 02-APR-1992

APPLICATION NUMBER: US 07/813,593

FILING DATE: 24-DEC-1991

APPLICATION NUMBER: US 07/793,065

FILING DATE: 15-NOV-1991

APPLICATION NUMBER: US 07/728,913

FILING DATE: 28-JUN-1991

APPLICATION NUMBER: US 07/679,666

FILING DATE: 02-APR-1991

ATTORNEY/AGENT INFORMATION:

NAME: Feit, Irving N.

REGISTRATION NUMBER: 28,601

REFERENCE/DOCKET NUMBER: LEM-3-7P

TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-645-1405

TELEFAX: 212-645-2054

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 3453 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: cDNA

HYPOTHETICAL: NO

ANTI-SENSE: NO

FRAGMENT TYPE: N-terminal

FEATURE:

NAME/KEY: mat_peptide

LOCATION: 112..3006

FEATURE:

NAME/KEY: sig_peptide

LOCATION: 31..111

FEATURE:

NAME/KEY: CDS

LOCATION: 31..3009

SEQUENCE DESCRIPTION: SEQ ID NO: 1:

US-09-919-408-1

Query Match 63.6%; Score 2227.2; DB 9; Length 3453;

Best Local Similarity 79.6%; Pred. No. 0;

Matches 2775; Conservative 0; Mismatches 653; Indels 56; Gaps 10;

Qy	35	GGGAGACCGCGCTCCGGAGGCGCATGCGCGGTGGC---GCGGACGCGGGCACCGTGC	91
Db	8	GGCTACCGCGCGCTCCGGAGGCGCATGCGCGGTGGCAGCGAGCGCGCGGCGTGC	67
Qy	92	CGCTGCTCGTGTGTTTCTCGATGATATTGGGACTATTACAAATCAAGATCTGCTG	151
Db	68	TGCTGCTGTGTTTGTGAGTAATGATTCGTGAGACGTTTACAAACCAAGACCTGCTG	127
Qy	152	TGATCAAGTGTGTTTAAATCAATCATAGAACAAATGATTCATCAGTGGGGAAGTCATCAT	211
Db	128	TGATCAAGTGTGTTTAAATCAATCATAGAACAAATGATTCATCAGTGGGGAAGTCATCAT	187
Qy	212	CATATCCCATGGTATCAGAAATCCCGGGAAGACCTCGGGTGTGCGTTGAGACCCAGAGCT	271

Db	188	CGTACCGAATGGTGGAGGATCCCGAAGAACCTCCAGTGTACCCCGAGCGCCAGAGTG	247
Qy	272	CAGGACAGTGTACGAAGCTGCGCTGTGGAAGTGTATCTGCTCCATCACACTGC	331
Db	248	AAGGACCGTATATGAAGCGGCCACCGTGGAGTGGCGGAGTCTCCATCACCTGC	307
Qy	332	AAGTGTGTCGATGCGCCAGGGAACATTTCTGTCTCTGGTCTTTAAGCACAGCTCCC	391
Db	308	AAGTGCAGCTCGCCACCCCGAGGAGCCTTCTCCCTCTGGGTCTTTAAGCACAGCTCCC	367
Qy	392	TGAATTGCCAGCCACATTTTGAATTTACAAACAGAGGAGTGTGTTTCCATGCTCAATTTGA	451
Db	368	TGGCTGCCAGCCGCACTTTGATTTACAAACAGAGGAATCGTTTCCATGCCATCTTGA	427
Qy	452	AAATGACAGAAACCCAAAGCTGGAGAACTACTTTTATTTTCAAGTGAAGCTACCAATT	511
Db	428	ACGTGACAGAGACCCAGGCGAGGAATACCTACTCCATATTTACAGAGGGAACGCGCAACT	487
Qy	512	ACACATATTTTACAGTGAATTAAGAAATACCTCTGCTTTACACATTAAGAGACCTT	571
Db	488	ACACAGTACTGTTTACAGTGAATTAAGAGATACACAGCTGTATGTGCTAAGGAGACCTT	547
Qy	572	ACTTTAGAAAAATGAAAAACCCAGGACCCCTGGTCTGCATATCTGAGAGCGTTCAGAGC	631
Db	548	ACTTTAGAAAGTGAAAAAACAGGATGCACTGCTCTGCATCTCCGAGGCTGTTCCGAGC	607
Qy	632	CGATCGTGAATGGGTGCTTTTGGGATTCACAGGGGGAAAGCTGTAAGAAAGAAAGTCCAG	691
Db	608	CCACTGTGGAGTGGTGTCTCTGCAGCTCCACAGGGAAGCTGTAAGAAAGAAAGCCCTG	667
Qy	692	CTGTTGTTAAAAAGGAGGAAAAAGTGTTCATGAATTTATTTGGGACGGACATAGGTGCT	751
Db	668	CTGTTGTGAGAAAGGAGGAAAAAGGTACTTTCATGAGTTGTTCCGAAACAGACATCAGATCT	727
Qy	752	GTGCCAGAAATGAACCTGGGCGAGGAATGCACAGCTGTTCCACATATAGATCTAATAACAA	811
Db	728	GTGCTAGAAATGCACTGGGCGCGAATGCACAAAGCTGTTCCCATAGATCTAATAACAGG	787
Qy	812	CTCCTCAGACACATTCGCCCAATTTATTTTAAAGTAGGGGAAACCCCTTATGGATAAGGT	871
Db	788	CTCCTCAGAGCACACTGCCCCAGTTTATTCCTGAAAGTGGGGGAAACCCCTTGTGGATCAGT	847
Qy	872	GCAAGCTGTTTATGTGAAACCATGGAATCGGGCTCACCTGGGGAATTAGAAACAAAGCAC	931
Db	848	GTAAGGCCATCCATGTGAACCATGGAATTCGGGCTCACCTGGGAGCTGGAAGACAAAGCCC	907
Qy	932	TCGAGGAGGGCAACTACTTTGAGATGAGTACTTCAACAAACAGACTATGATACGGA	991
Db	908	TGGAGGAGGGCAGCTACTTTGAGATGAGTACTTCAACAAACAGGACCATGATTCGGA	967
Qy	992	TTCTGTTTGTCTTTGTATCATCAGTGGCAAGAAACGACACCGGATATCTACACTTGTCTCT	1051
Db	968	TTCTTGTGGCTTTGTGCTCTTCGTGGGAGGAGGACACCGGATATTACACCTGCTCTT	1027
Qy	1052	CTTCAAAAGCATCCAGTCAATCAGCTTTGGTTTACCTGTGAGGAAGGATTTATAAATG	1111
Db	1028	CCTCAAAAGCACCCAGCGCAGTCACTGTTGGTGGACCATCTAGAAAAAAGGGTTTATAAAG	1087
Qy	1112	CTACCAATTCAGTGAAGATTTGAAATTCGCAATATGAGAGCTTTGTTTCTCTGTCATCA	1171
Db	1088	CTACCACTCGCAAGAGAGATGAAATTCGCAATATGAGAGCTTTGTTTCTCTGTCATCA	1147
Qy	1172	GGTTTAAAGCCTACCCCAATCAGATGTAGTGGACCTCTCTCGAAAAATCATTTCTCTT	1231
Db	1148	GGTTTAAAGCGTACCCAGAAATCCGATGCGAGTGGATCTCTCTCAAGCCTCATTTCTCTT	1207
Qy	1232	GTGAGCAAAAGGGTCTTTGATAACGGATACAGCATATCCAAAGTTTGTCAATCATAGCAC	1291
Db	1208	GTGAACAGAGAGGCTCGAGGATGGGTACAGCATATCTAAATTTTGGCATCATAGAACA	1267
Qy	1292	AGCCAGAGAGATATATTTTCCATCCAGAAAAATGATGATGCCCAATTTACCAAAATGTTCA	1351
Db	1268	AGCCAGGAGAGTACATATTTCTATGCAAAAAATGATGAGCCCGCTTCAACCAAAATGTTCA	1327

Db 3450 AAAA 3453

|||||

RESULT 5

US-09-872-136-1

; Sequence 1, Application US/09872136

; Patent No. US20020119545A1

; GENERAL INFORMATION:

APPLICANT: Lemischka, Ihor R.

TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL

NUMBER OF SEQUENCES: 10

CORRESPONDENCE ADDRESS:

ADDRESSEE: Imclone Systems Incorporated

STREET: 180 Varick Street

CITY: New York

STATE: New York

COUNTRY: U.S.A.

ZIP: 10014

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/872,136

FILING DATE: 01-Jun-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/09/208,786

FILING DATE: <Unknown>

APPLICATION NUMBER: US/09/021,324

FILING DATE: <Unknown>

APPLICATION NUMBER: US/07/977,451

FILING DATE: 1992-11-19

APPLICATION NUMBER: US 07/906,397

FILING DATE: 26-JUN-1992

APPLICATION NUMBER: US PCT/US92/05401

FILING DATE: 26-JUN-1992

APPLICATION NUMBER: TW 81102961

FILING DATE: 15-APR-1992

APPLICATION NUMBER: US PCT/US92/02750

FILING DATE: 02-APR-1992

APPLICATION NUMBER: US 07/813,593

FILING DATE: 24-DEC-1991

APPLICATION NUMBER: US 07/793,065

FILING DATE: 15-NOV-1991

APPLICATION NUMBER: US 07/728,913

FILING DATE: 28-JUN-1991

APPLICATION NUMBER: US 07/679,666

FILING DATE: 02-APR-1991

ATTORNEY/AGENT INFORMATION:

NAME: Feit, Irving N.

REGISTRATION NUMBER: 28,601

REFERENCE/DOCKET NUMBER: LEM-3-7P

TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-645-1405

TELEFAX: 212-645-2054

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 3453 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: cDNA

HYPOTHETICAL: NO

ANTI-SENSE: NO

FRAGMENT TYPE: N-terminal

FEATURE:

NAME/KEY: mat_peptide

LOCATION: 112..3006

FEATURE:

NAME/KEY: sig_peptide

LOCATION: 31..111

FEATURE:

NAME/KEY: CDS

LOCATION: 31..3009

SEQUENCE DESCRIPTION: SEQ ID NO: 1:

US-09-872-136-1

Query Match 63.6%; Score 2227.2; DB 10; Length 3453;

Best Local Similarity 79.6%; Pred. No. 0;

Matches 2775; Conservative 0; Mismatches 653; Indels 56; Gaps 10;

QY 35 GGGACACCCCGGCTCCGAGGCGCATCCGGGCTGGC---GGCGACCGGCGCACCGTGC 91

Db 8 GGCTACCGCGCGCTCCGAGGCGCATCCGGGCTGGCAGCGAGCGAGCGCGGCTGC 67

QY 92 CGCTGCTGTTGTTTCTGCAATGATATTGGGACTATTACAAATCAAGATCTGCCTG 151

Db 68 TGCTGCTGTTGTTTGTCTGCAATGATATTGGGACTATTACAAATCAAGATCTGCCTG 127

QY 152 TGATCAAGTGTGTTTAAATCAATCATAGAAATGATTCATCAGTGGGAAAGTCAAT 211

Db 128 TGATCAAGTGTGTTTAAATCAATCATAGAAATGATTCATCAGTGGGAAAGTCAAT 187

QY 212 CATATCCATGTTATCAGATCCCGGAGAGCTCCGGTGTGCTTGGACCCAGAGCT 271

Db 188 CGTACCGAATGTCGAGAGTCCCGAGAGACCTCCAGTGTACCCCGAGGCGCGAGAGT 247

QY 272 CAGGACAGAGTGTACGAGCTCCGCTGTGGAAGTGTGATGATCTCTCTCCATCACATCG 331

Db 248 AAGGACCGTATATGAGCGGCCACCGTGGAGTGGCGAGTCTGGTTCATCACCTGC 307

QY 332 AAGTGTGTCGATGCCCGAGGAAATTTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCT 391

Db 308 AAGTGTGTCGATGCCCGAGGAAATTTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCT 367

QY 392 TGAATGCCAGCCACATTTGATTTAATAAGAGAGGAGTGTGTTCCATGTCATTTGA 451

Db 368 TGGGCTGCCAGCGCCTTTGATTTTACAAAACAGAGGAATCGTTTCCATGGCCATCTTGA 427

QY 452 AAATCAGAGAAACCCAGTGTGGAATACCTATTTTATTCAGAGTGAAGTACCAAT 511

Db 428 ACGTGACAGAGACCCAGGAGAGATACCTACTCTATTCAGAGCGGAAACCGCGCACT 487

QY 512 ACACAAATTTGTTACAGTGTATAGAAATACCCCTGCTTTTACACATTAAGAGACCTT 571

Db 488 ACACAGTACTGTTACAGTGTATGTAAGAGATACACAGCTGTATGTGTAAGAGACCTT 547

QY 572 ACTTTAGAAAATGGAACCCAGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 631

Db 548 ACTTTAGGAAGATGGAACCCAGAGTGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 607

QY 632 CGATCGTGAATGGTGTGCTTTCGATTCACAGGGGAAAGCTGTAAGAGAAATCCAG 691

Db 608 CCATGTGGAGTGGTGTCTCTGAGCTCCCAAGGAAAGCTGTAAGAGAAAGGCGCTG 667

QY 692 CTGTTGTTTAAAGAGGAAAGTGTCTTCAATGATTTTGGGACGACATAGGTGCT 751

Db 668 CTGTTGTCAGAAAGGAGGAAAGTGTCTTCAATGATTTTGGGACGACATAGGTGCT 727

QY 752 GTGCCAGAAATGAAGTGGGAGGAAATGCAACAGGCTGTTCAATAGATCTAAATCAA 811

Db 728 GTGCTAGAAATGCACTGGGCGCGAAATGCAACAGCTGTTCAACATAGATCTAAACAG 787

QY 812 CTCTCAGACCATGTCACCAATTTTCTTAAAGTGGGAAAGCTTATGATAGGT 871

Db 788 CTCTCAGAGCACATGCCCCAGTATTCTTCTGAAAGTGGGAAAGCTTGTGATGAT 847

QY 872 GCAAAGCTGTTTCATGTGAACCATGATTCGGGCTCACCTCGGAAATTTAGAAAACAAAGCAC 931

Db 848 GTAAGGCCATCCATGTGAACCATGATTCGGGCTCACCTCGGAGCTGGAGACAAAGCCC 907

QY 932 TCGAGGAGGCGCAACTACTTTGAGATGAGTACCTATTCAACAAACAGAACTATGATACGGA 991

QY 3146 TGCTGCTTACCAGACTTTTCTCTAGAAGCCGCTGCGTTTACTCTGTGTTTCAAAGGA 3205
Db 3097 CGTTGCTTCCGCTGGACTTTTCTCTAGATGCTGTCTGCCATTACTC-----CAAAGTGA 3149
QY 3206 CTTTGTGTAATCAATCATCTCTGCACAGGAGGAGCTGATATGAACCTTTATG 3265
Db 3150 CTTCTATAAATCAAACCTCTCTCCACAGCGGGAGAGCCAAATAATGAGACTTGTGG 3209
QY 3266 GAGCATTGATCTGCATCAAGGSCCTTCTCAGCGCGCTTGAGTGAATTGTGTACCTGAAG 3325
Db 3210 TGAGCCGCTACCTCGGGGCTTTCACGAGCTTGAGGGGAAGCCATGTATCTGAA 3269
QY 3326 TACAGTATATCTTGTAAATACATAAAACAAA-----AGCATTGTTGAAGGAGAGC 3378
Db 3270 TATAGTATATCTTGTAAATACGTGAACCAACCAACCCGTTTTTGTGTAAGGGAAGC 3329
QY 3379 TAATATGATTTTT--AGTCTATGTTTAAATAATATGTAATTTTTTTCAGCTATTAG 3436
Db 3330 TAAATATGATTTTTTAAATAATCTATGTTTTTAAATAATCTATGTAACTTTTTTCATCTATTAG 3389
QY 3437 TGATATATTTTATGCTGGAATAAAATTTCTACTACAGAAAAAATAAATAAATAA 3496
Db 3390 TGATATATTTTATGATGAATAAACTTTCTACTGTAATAAATAAATAAATAAATAA 3449
QY 3497 AAAA 3500
Db 3450 AAAA 3453

RESULT 6
US-09-796-692-7538
; Sequence 7538, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
; FILE REFERENCE: 2077.001200
; CURRENT APPLICATION NUMBER: US/09/796,692
; PRIOR FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO: 7538
; LENGTH: 645
; TYPE: DNA
; ORGANISM: Homo sapiens

US-09-796-692-7538

Query Match 18.1%; Score 633; DB 10; Length 645;
Best Local Similarity 99.8%; Pred. No. 3.6e-150;
Matches 644; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 2059 CTGGAGCCACGAGAAATATTGTGAACCTGTGGGGCGTGACACTGTTCAGAGCAAT 2118
Db 1 CTGGAGCCACGAGAAATATTGTGAACCTGTGGGGCGTGACACTGTTCAGAGCAAT 60
QY 2119 TACTTGAATTTTGAATCTGTGTGATCTTCTCAACTATCTTAAGAAGTAAAGA 2178
Db 61 TACTTGAATTTTGAATCTGTGTGATCTTCTCAACTATCTTAAGAAGTAAAGA 120
QY 2179 GAAAAATTTTCAAGGACTTGGACAGAGATTTTCAAGGAAACAAATTTTACCC 2238
Db 121 GAAAAATTTTCAAGGACTTGGACAGAGATTTTCAAGGAAACAAATTTTACCC 180
QY 2239 ACTTTCCAAATCACAATTCAGCATGCTGCTTCAAGAGAGTTTCAAGAGTACACCCG 2298
Db 181 ACTTTCCAAATCACAATTCAGCATGCTGCTTCAAGAGAGTTTCAAGAGTACACCCG 240
QY 2299 GACTCGGATCAAAATCTCAGGGCTTTCATGGGAATTCATTTCACTCTGAAGTAAATTTGAA 2358
Db 241 GACTCGGATCAAAATCTCAGGGCTTTCATGGGAATTCATTTCACTCTGAAGTAAATTTGAA 300
QY 2359 TATGAAAAACCAAAAGGCTGGAAGAAGGAGGAGTCTTGAATGTCTTCACTTTGAAGAT 2418
Db 301 TATGAAAAACCAAAAGGCTGGAAGAAGGAGGAGTCTTGAATGTCTTCACTTTGAAGAT 360
QY 2419 CTTCTTGTGCTTGCATATCAAGTTGCCAAAGAAATGGAATTTCTGGAATTTAAGTCGTGT 2478
Db 361 CTTCTTGTGCTTGCATATCAAGTTGCCAAAGAAATGGAATTTCTGGAATTTAAGTCGTGT 420
QY 2479 GTTTCACAGAGACTGCGCCAGGAAAGCTGTTGTCCACCACGGGAAAGTGTGAAGATA 2538
Db 421 GTTTCACAGAGACTGCGCCAGGAAAGCTGTTGTCCACCACGGGAAAGTGTGAAGATA 480
QY 2539 TGTGACTTTGGATTGGCTCGAGATATCATGAGTGAATTCCAACTATGTTTGTAGGGGCAAT 2598
Db 481 TGTGACTTTGGATTGGCTCGAGATATCATGAGTGAATTCCAACTATGTTTGTAGGGGCAAT 540
QY 2599 GCCCGTCTGCTGTAAATGGATGGCCCGGAAAGGCTG-TTTGAAGGATCTTACACCAT 2657
Db 541 GCCCGTCTGCTGTAAATGGATGGCCCGGAAAGGCTG-TTTGAAGGATCTTACACCAT 600
QY 2658 TAAGAGTGAATGCTGCTCATATGGAATATTACTGTGGAAATCTT 2702
Db 601 TAAGAGTGAATGCTGCTCATATGGAATATTACTGTGGAAATCTT 645

RESULT 7
US-10-040-862-7538
; Sequence 7538, Application US/10040862
; Publication No. US20030078396A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Thera
; FILE REFERENCE: 014058-013520US
; CURRENT APPLICATION NUMBER: US/10/040,862
; CURRENT FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303

;; PRIOR FILING DATE: 2000-04-28
;; PRIOR APPLICATION NUMBER: US 60/200,779
;; PRIOR FILING DATE: 2000-04-28
;; PRIOR APPLICATION NUMBER: US 60/200,999
;; PRIOR FILING DATE: 2000-05-01
;; PRIOR APPLICATION NUMBER: US 60/202,084
;; PRIOR FILING DATE: 2000-05-04
;; PRIOR APPLICATION NUMBER: US 60/206,201
;; PRIOR FILING DATE: 2000-05-22
;; PRIOR APPLICATION NUMBER: US 60/218,950
;; PRIOR FILING DATE: 2000-07-14
;; PRIOR APPLICATION NUMBER: US 60/222,903
;; PRIOR FILING DATE: 2000-08-03
;; PRIOR APPLICATION NUMBER: US 60/223,416
;; PRIOR FILING DATE: 2000-08-04
;; PRIOR APPLICATION NUMBER: US 60/223,378
;; PRIOR FILING DATE: 2000-08-07
;; PRIOR APPLICATION NUMBER: US 09/796,692
;; PRIOR FILING DATE: 2001-03-01
;; NUMBER OF SEQ ID NOS: 10467
;; SOFTWARE: FastSeq for Windows Version 3.0
;; SEQ ID NO 7538
;; LENGTH: 645
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-10-040-862-7538

Query Match
Best Local Similarity 18.1%; Score 633; DB 14; Length 645;
Matches 644; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 2059 CTGGGAAGCCAGAGATATTGTAACCTGCTGGGGGGTGCACACTGTCAGGACCAATT 2118
DB 1 CTGGGAAGCCAGAGATATTGTAACCTGCTGGGGGGTGCACACTGTCAGGACCAATT 60
QY 2119 TACTTGATTTTGAATCTGCTATGGTGATCTTCTCAACTATCTAAGAAGTAAAGA 2178
DB 61 TACTTGATTTTGAATCTGCTATGGTGATCTTCTCAACTATCTAAGAAGTAAAGA 120
QY 2179 GAAAAATTCAGACACTTGGACAGAGATTTTCAAGGAACACAATTTTACCCCC 2238
DB 121 GAAAAATTCAGACACTTGGACAGAGATTTTCAAGGAACACAATTTTACCCCC 180
QY 2239 ACTTTCCAAATCACATCCAAATTCAGCATGCTGTTCAAGGAAGTTCAGATACACCCG 2298
DB 181 ACTTTCCAAATCACATCCAAATTCAGCATGCTGTTCAAGGAAGTTCAGATACACCCG 240
QY 2299 GACTCGGATCAAAATCTCAGGCTTCATCGGAATTCATTTCTCTGAAGATGAATTTGAA 2358
DB 541 GACTCGGATCAAAATCTCAGGCTTCATCGGAATTCATTTCTCTGAAGATGAATTTGAA 600
QY 2658 TAAGAGTGATCTGCTGGTTCATATGGAATATTACTGTGGGAATCTT 2702

DB 601 TAAGAGTGATGCTGCTCATATGCAATATTACTGTGGAAATCTT 645
RESULT 8
US-09-796-692-8085
;; Sequence 8085, Application US/09796692
;; Publication No. US20020198362A1
;; GENERAL INFORMATION:
;; APPLICANT: Gaiger, Alexander
;; APPLICANT: Algate, Paul A.
;; APPLICANT: Mannion, Jane
;; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERA
;; FILE OF INVENTION: HEMATOLOGICAL MALIGNANCIES
;; FILE REFERENCE: 2077,001200
;; CURRENT APPLICATION NUMBER: US/09/796,692
;; CURRENT FILING DATE: 2001-03-01
;; PRIOR APPLICATION NUMBER: 60/186,126
;; PRIOR FILING DATE: 2000-03-01
;; PRIOR APPLICATION NUMBER: 60/190,479
;; PRIOR FILING DATE: 2000-03-17
;; PRIOR APPLICATION NUMBER: 60/200,545
;; PRIOR FILING DATE: 2000-04-27
;; PRIOR APPLICATION NUMBER: 60/200,303
;; PRIOR FILING DATE: 2000-04-28
;; PRIOR APPLICATION NUMBER: 60/200,779
;; PRIOR FILING DATE: 2000-04-28
;; PRIOR APPLICATION NUMBER: 60/206,201
;; PRIOR FILING DATE: 2000-05-22
;; PRIOR APPLICATION NUMBER: 60/218,950
;; PRIOR FILING DATE: 2000-07-14
;; PRIOR APPLICATION NUMBER: 60/222,903
;; PRIOR FILING DATE: 2000-08-03
;; PRIOR APPLICATION NUMBER: 60/223,416
;; PRIOR FILING DATE: 2000-08-04
;; PRIOR APPLICATION NUMBER: 60/223,378
;; NUMBER OF SEQ ID NOS: 9597
;; SOFTWARE: FastSeq for Windows Version 3.0
;; SEQ ID NO 8085
;; LENGTH: 610
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-09-796-692-8085

Query Match
Best Local Similarity 17.4%; Score 608.4; DB 10; Length 610;
Matches 609; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2059 CTGGGAAGCCAGAGATATTGTAACCTGCTGGGGGGTGCACACTGTCAGGACCAATT 2118
DB 1 CTGGGAAGCCAGAGATATTGTAACCTGCTGGGGGGTGCACACTGTCAGGACCAATT 60
QY 2119 TACTTGATTTTGAATCTGCTATGGTGATCTTCTCAACTATCTAAGAAGTAAAGA 2178
DB 61 TACTTGATTTTGAATCTGCTATGGTGATCTTCTCAACTATCTAAGAAGTAAAGA 120
QY 2179 GAAAAATTCAGACACTTGGACAGAGATTTTCAAGGAACACAATTTTACCCCC 2238
DB 121 GAAAAATTCAGACACTTGGACAGAGATTTTCAAGGAACACAATTTTACCCCC 180
QY 2239 ACTTTCCAAATCACATCCAAATTCAGCATGCTGTTCAAGGAAGTTCAGATACACCCG 2298
DB 181 ACTTTCCAAATCACATCCAAATTCAGCATGCTGTTCAAGGAAGTTCAGATACACCCG 240
QY 2299 GACTCGGATCAAAATCTCAGGCTTCATCGGAATTCATTTCTCTGAAGATGAATTTGAA 2358
DB 241 GACTCGGATCAAAATCTCAGGCTTCATCGGAATTCATTTCTCTGAAGATGAATTTGAA 300
QY 2359 TATGAAAAACCAAAAGGCTCGGAAGAGAGGAGGACTTGAATGTGCTTACATTTGAAGAT 2418

;; PRIOR FILING DATE: 2000-05-01
;; PRIOR APPLICATION NUMBER: 60/202,084
;; PRIOR FILING DATE: 2000-05-04
;; PRIOR APPLICATION NUMBER: 60/206,201
;; PRIOR FILING DATE: 2000-05-22
;; PRIOR APPLICATION NUMBER: 60/218,950
;; PRIOR FILING DATE: 2000-07-14
;; PRIOR APPLICATION NUMBER: 60/222,903
;; PRIOR FILING DATE: 2000-08-03
;; PRIOR APPLICATION NUMBER: 60/223,416
;; PRIOR FILING DATE: 2000-08-04
;; PRIOR APPLICATION NUMBER: 60/223,378
;; PRIOR FILING DATE: 2000-08-07
;; NUMBER OF SEQ ID NOS: 9597
;; SOFTWARE: FastSeq for Windows Version 3.0
;; SEQ ID NO 8132
;; LENGTH: 608
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-10-040-862-8132

Query Match
Best Local Similarity 17.3%; Score 606.4; DB 10; Length 608;
Matches 607; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2288 AGATACACCGGACTCGGATCAAACTCTCAGGCTTCATCGGAATTCATTTCCTCTGAAG 2347
DB 608 AGATACACCGGACTCGGATCAAACTCTCAGGCTTCATCGGAATTCATTTCCTCTGAAG 549
QY 2348 ATGAAATTGAATATGAAACCCAAAGAGGCTGGAAGAGGAGGAGGACTTGAATGTGCTTA 2407
DB 548 ATGAAATTGAATATGAAACCCAAAGAGGCTGGAAGAGGAGGAGGACTTGAATGTGCTTA 489
QY 2408 CATTGGAAGATCTCTTTGCTTTGCTATCAAGTTGCGCAAGGAATGGAATTTCTGGAAT 2467
DB 488 CATTGGAAGATCTCTTTGCTTTGCTATCAAGTTGCGCAAGGAATGGAATTTCTGGAAT 429
QY 2468 TTAAGTCGTGTGTTACAGAGACCTGGCCGCGAGAACGCTGTTGTACCCACCGGGAAG 2527
DB 428 TTAAGTCGTGTGTTACAGAGACCTGGCCGCGAGAACGCTGTTGTACCCACCGGGAAG 369
QY 2528 TGGTGAAGATATGTGACTTTGGATTGGCTCGAGATATCATGAGTGATTCACCACTATGTTG 2587
DB 368 TGGTGAAGATATGTGACTTTGGATTGGCTCGAGATATCATGAGTGATTCACCACTATGTTG 309
QY 2588 TCAGGGCAATGCCCGTCTGCTGTAATAATGGATGGCCCGGAAAGCCTGTTTGAAGCA 2647
DB 308 TCAGGGCAATGCCCGTCTGCTGTAATAATGGATGGCCCGGAAAGCCTGTTTGAAGCA 249

;; Publication No. US20030078396A1
;; GENERAL INFORMATION:
;; APPLICANT: Gaiger, Alexander
;; APPLICANT: Algate, Paul A.
;; APPLICANT: Mannion, Jane
;; APPLICANT: Retter, Marc
;; APPLICANT: Corixa Corporation
;; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and There
;; FILE REFERENCE: 014058-013520US
;; CURRENT APPLICATION NUMBER: US/10/040,862
;; CURRENT FILING DATE: 2001-11-06
;; PRIOR APPLICATION NUMBER: US 60/186,126
;; PRIOR FILING DATE: 2000-03-01
;; PRIOR APPLICATION NUMBER: US 60/190,479
;; PRIOR FILING DATE: 2000-03-17
;; PRIOR APPLICATION NUMBER: US 60/200,545
;; PRIOR FILING DATE: 2000-04-27
;; PRIOR APPLICATION NUMBER: US 60/200,303
;; PRIOR FILING DATE: 2000-04-28
;; PRIOR APPLICATION NUMBER: US 60/200,779
;; PRIOR FILING DATE: 2000-04-28
;; PRIOR APPLICATION NUMBER: US 60/206,201
;; PRIOR FILING DATE: 2000-05-22
;; PRIOR APPLICATION NUMBER: US 60/202,084
;; PRIOR FILING DATE: 2000-05-04
;; PRIOR APPLICATION NUMBER: US 60/218,950
;; PRIOR FILING DATE: 2000-07-14
;; PRIOR APPLICATION NUMBER: US 60/222,903
;; PRIOR FILING DATE: 2000-08-03
;; PRIOR APPLICATION NUMBER: US 60/223,416
;; PRIOR FILING DATE: 2000-08-04
;; PRIOR APPLICATION NUMBER: US 60/223,378
;; PRIOR FILING DATE: 2000-08-07
;; PRIOR APPLICATION NUMBER: US 09/796,692
;; PRIOR FILING DATE: 2001-03-01
;; NUMBER OF SEQ ID NOS: 10467
;; SOFTWARE: FastSeq for Windows Version 3.0
;; SEQ ID NO 8132
;; LENGTH: 608
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-10-040-862-8132

Query Match
Best Local Similarity 17.3%; Score 606.4; DB 14; Length 608;
Matches 607; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2288 AGATACACCGGACTCGGATCAAACTCTCAGGCTTCATCGGAATTCATTTCCTCTGAAG 2347
DB 608 AGATACACCGGACTCGGATCAAACTCTCAGGCTTCATCGGAATTCATTTCCTCTGAAG 549
QY 2348 ATGAAATTGAATATGAAACCCAAAGAGGCTGGAAGAGGAGGAGGACTTGAATGTGCTTA 2407
DB 548 ATGAAATTGAATATGAAACCCAAAGAGGCTGGAAGAGGAGGAGGACTTGAATGTGCTTA 489
QY 2408 CATTGGAAGATCTCTTTGCTTTGCTATCAAGTTGCGCAAGGAATGGAATTTCTGGAAT 2467
DB 488 CATTGGAAGATCTCTTTGCTTTGCTATCAAGTTGCGCAAGGAATGGAATTTCTGGAAT 429
QY 2468 TTAAGTCGTGTGTTACAGAGACCTGGCCGCGAGAACGCTGTTGTACCCACCGGGAAG 2527
DB 428 TTAAGTCGTGTGTTACAGAGACCTGGCCGCGAGAACGCTGTTGTACCCACCGGGAAG 369
QY 2528 TGGTGAAGATATGTGACTTTGGATTGGCTCGAGATATCATGAGTGATTCACCACTATGTTG 2587
DB 368 TGGTGAAGATATGTGACTTTGGATTGGCTCGAGATATCATGAGTGATTCACCACTATGTTG 309
QY 2588 TCAGGGCAATGCCCGTCTGCTGTAATAATGGATGGCCCGGAAAGCCTGTTTGAAGCA 2647
DB 308 TCAGGGCAATGCCCGTCTGCTGTAATAATGGATGGCCCGGAAAGCCTGTTTGAAGCA 249

QY 2648 TCTACACCAATTAAAGTAGTCTGGTCATATGGAATATTACTGTGGGAAATCTTCTCAC 2707
DB 248 TCTACACCAATTAAAGTAGTCTGGTCATATGGAATATTACTGTGGGAAATCTTCTCAC 189
QY 2708 TTGGTGTGAATCCCTACCTGGCATTCCGGTTGATGCTTAACCTTCTACAACTGATTGAAA 2767
DB 188 TTGGTGTGAATCTTTACCTGGCATTCCGGTTGATGCTTAACCTTCTACAACTGATTGAAA 129
QY 2768 ATGATTTAAATGGATCAGCCATTTTATGCTACAGAGAATAATACATTAATATGCAAT 2827
DB 128 ATGATTTAAATGGATCAGCCATTTTATGCTACAGAGAATAATACATTAATATGCAAT 69
QY 2828 CCTGCTGGGCTTTTGACTCAAGGAAACGGCCATCCTTCCCTAAATTTGACTTCGTTTTTGA 2887
DB 68 CCTGCTGGGCTTTTGACTCAAGGAAACGGCCATCCTTCCCTAAATTTGACTTCGTTTTTGA 9
QY 2888 GATGTCAG 2895
DB 8 GATGTCAG 1

RESULT 12
US-09-796-692-8927
; Sequence 8927, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
; FILE REFERENCE: 2077.001200
; CURRENT APPLICATION NUMBER: US/09/796,692
; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8927
; LENGTH: 605
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-796-692-8927
Query Match 17.3%; Score 605; DB 10; Length 605;
Best Local Similarity 100.0%; Pred. No. 4.5e-143;
Matches 605; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2059 CTGGGAAGCCAGAGAAATTGTGAACTTGTGGGGGCTGCACACTGTTCAGGACCAATT 2118
,|||||

DB 1 CTGGGAAGCCAGAGAAATTGTGAACTTGTGGGGGCTGCACACTGTTCAGGACCAATT 60
QY 2119 TACTTGAATTTTGAATACTGTGTCTATGCTGATCTTCTCAACTATCTTAAGAAAGTAAAGA 2178
DB 61 TACTTGAATTTTGAATACTGTGTCTATGCTGATCTTCTCAACTATCTTAAGAAAGTAAAGA 120
QY 2179 GAAAAATTTTACAGAGCTTGGACAGAGATTTTCAAGNACACAAATTTTCAAGTTTACCCC 2238
DB 121 GAAAAATTTTACAGAGCTTGGACAGAGATTTTCAAGNACACAAATTTTCAAGTTTACCCC 180
QY 2239 ACTTTTCAATCACATCCAAATTTCCAGCATGCTGTTTCAAGAGAAAGTTTCAGATACACCCG 2298
DB 181 ACTTTTCAATCACATCCAAATTTCCAGCATGCTGTTTCAAGAGAAAGTTTCAGATACACCCG 240
QY 2299 GACTCGGATCAAAATCTCAGGGCTTTCATGGGAATTCATTTCATCTGAAGATGAAATGAA 2358
DB 241 GACTCGGATCAAAATCTCAGGGCTTTCATGGGAATTCATTTCATCTGAAGATGAAATGAA 300
QY 2359 TATGAAAACCAAAAAAGGCTGGAAAGAGAGAGAGACTTGAATGCTTACATTTGAAGAT 2418
DB 301 TATGAAAACCAAAAAAGGCTGGAAAGAGAGAGAGACTTGAATGCTTACATTTGAAGAT 360
QY 2419 CTCTTTGCTTTTGCATATCAAGTTTGCACAAAGAAATGGAATTTCTGGAATTTTAAAGTCGTG 2478
DB 361 CTCTTTGCTTTTGCATATCAAGTTTGCACAAAGAAATGGAATTTCTGGAATTTTAAAGTCGTG 420
QY 2479 GTTCACAGAGACCTGGCCGCCAGGAACTGCTTGTCAACCCAGGAAAGTGTGAGAGATA 2538
DB 421 GTTCACAGAGACCTGGCCGCCAGGAACTGCTTGTCAACCCAGGAAAGTGTGAGAGATA 480
QY 2539 TGTGACTTTGGAATTTGGCTCGAGATATCATAGTAGTATTCACACTATGTTGTCTAGGGGCAAT 2598
DB 481 TGTGACTTTGGAATTTGGCTCGAGATATCATAGTAGTATTCACACTATGTTGTCTAGGGGCAAT 540
QY 2599 GCCGCTCTGCTGTAAATGATGCTGCCCCCAAGAGCTGTTTGAAGGCACTCTACACCAATT 2658
DB 541 GCCGCTCTGCTGTAAATGATGCTGCCCCCAAGAGCTGTTTGAAGGCACTCTACACCAATT 600
QY 2659 AAGAG 2663
DB 601 AAGAG 605
RESULT 13
US-10-040-862-8927
; Sequence 8927, Application US/10040862
; Publication No. US20030078396A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therap
; FILE REFERENCE: 014058-013520US
; CURRENT APPLICATION NUMBER: US/10/040,862
; CURRENT FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201

; PRIOR FILING DATE: 2000-05-22
 ; PRIOR APPLICATION NUMBER: US 60/218,950
 ; PRIOR FILING DATE: 2000-07-14
 ; PRIOR APPLICATION NUMBER: US 60/222,903
 ; PRIOR FILING DATE: 2000-08-03
 ; PRIOR APPLICATION NUMBER: US 60/223,416
 ; PRIOR FILING DATE: 2000-08-04
 ; PRIOR APPLICATION NUMBER: US 60/223,378
 ; PRIOR FILING DATE: 2000-08-07
 ; PRIOR APPLICATION NUMBER: US 09/796,692
 ; PRIOR FILING DATE: 2001-03-01
 ; NUMBER OF SEQ ID NOS: 10467
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 8927
 ; LENGTH: 605
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-10-040-862-8927

Query Match
 Best Local Similarity 100.0%; Score 605; DB 14; Length 605;
 Matches 605; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	2059	CTGGAAGCCACGAGATATTTGAACTGCTGGGGCGTGACACTGTGACAGCAATTT	2118
Db	1	CTGGAAGCCACGAGATATTTGAACTGCTGGGGCGTGACACTGTGACAGCAATTT	60
Qy	2119	TACTTGATTTTGAATCTGTTGCTATGCTGATCTTCAACTATCTAAGAGTAAAGA	2178
Db	61	TACTTGATTTTGAATCTGTTGCTATGCTGATCTTCAACTATCTAAGAGTAAAGA	120
Qy	2179	GAAAAATTTTCAGGACTTTGGACAGAGATTTTCAAGGAACACAATTTTACCC	2238
Db	121	GAAAAATTTTCAGGACTTTGGACAGAGATTTTCAAGGAACACAATTTTACCC	180
Qy	2239	ACTTCCAAATCACATCCAAATTCAGCATGCTGTTCAAGAGAGATTTGACATACCCG	2298
Db	181	ACTTCCAAATCACATCCAAATTCAGCATGCTGTTCAAGAGAGATTTGACATACCCG	240
Qy	2299	GACTCGGATCAAAATCTCAGGGCTTCATGGGAATTCATTTCACTCTGAAGATGAA	2358
Db	241	GACTCGGATCAAAATCTCAGGGCTTCATGGGAATTCATTTCACTCTGAAGATGAA	300
Qy	2359	TATGAAAAACCAAAAAGGCTTGAAGAGAGAGAGAGACTTGAATGCTTACATTTGA	2418
Db	301	TATGAAAAACCAAAAAGGCTTGAAGAGAGAGAGAGACTTGAATGCTTACATTTGA	360
Qy	2419	CTTCTTTGCTTTGCATATCAAGTTGCCAAAGGAATGGAATTTCTGGAATTTAAGTCGT	2478
Db	361	CTTCTTTGCTTTGCATATCAAGTTGCCAAAGGAATGGAATTTCTGGAATTTAAGTCGT	420
Qy	2479	GTTCCAGAGACCTGGCCGACAGAACGTTGTCACCCACGGGAAAGTGGTGAAGATA	2538
Db	421	GTTCCAGAGACCTGGCCGACAGAACGTTGTCACCCACGGGAAAGTGGTGAAGATA	480
Qy	2539	TGTGACTTTTGGATTGGCTCGAGATATCATAGTGAATTCCTCAACTATGTTGTCAGGGCAAT	2598
Db	481	TGTGACTTTTGGATTGGCTCGAGATATCATAGTGAATTCCTCAACTATGTTGTCAGGGCAAT	540
Qy	2599	GCCCGTCTGCCGTAAATGATGGCCGCCGAAAGCCCTGTTTGAAGGATCTACACCAAT	2658
Db	541	GCCCGTCTGCCGTAAATGATGGCCGCCGAAAGCCCTGTTTGAAGGATCTACACCAAT	600
Qy	2659	AAGAG 2663	
Db	601	AAGAG 605	

RESULT 14
 US-09-796-692-8273
 ; Sequence 8273, Application US/09796692
 ; Publication No. US20020198362A1
 ; GENERAL INFORMATION:

; APPLICANT: Gaiger, Alexander
 ; APPLICANT: Algate, Paul A.
 ; APPLICANT: Mannion, Jane
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERA
 ; FILE REFERENCE: 2077.001200
 ; CURRENT APPLICATION NUMBER: US/09/796,692
 ; CURRENT FILING DATE: 2001-03-01
 ; PRIOR APPLICATION NUMBER: 60/186,126
 ; PRIOR FILING DATE: 2000-03-01
 ; PRIOR APPLICATION NUMBER: 60/190,479
 ; PRIOR FILING DATE: 2000-03-17
 ; PRIOR APPLICATION NUMBER: 60/200,545
 ; PRIOR FILING DATE: 2000-04-27
 ; PRIOR APPLICATION NUMBER: 60/200,303
 ; PRIOR FILING DATE: 2000-04-28
 ; PRIOR APPLICATION NUMBER: 60/200,779
 ; PRIOR FILING DATE: 2000-04-28
 ; PRIOR APPLICATION NUMBER: 60/200,999
 ; PRIOR FILING DATE: 2000-05-01
 ; PRIOR APPLICATION NUMBER: 60/202,084
 ; PRIOR FILING DATE: 2000-05-04
 ; PRIOR APPLICATION NUMBER: 60/206,201
 ; PRIOR FILING DATE: 2000-05-22
 ; PRIOR APPLICATION NUMBER: 60/218,950
 ; PRIOR FILING DATE: 2000-07-14
 ; PRIOR APPLICATION NUMBER: 60/222,903
 ; PRIOR FILING DATE: 2000-08-03
 ; PRIOR APPLICATION NUMBER: 60/223,416
 ; PRIOR FILING DATE: 2000-08-04
 ; PRIOR APPLICATION NUMBER: 60/223,378
 ; PRIOR FILING DATE: 2000-08-07
 ; NUMBER OF SEQ ID NOS: 9597
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 8273
 ; LENGTH: 601
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURES:
 ; NAME/KEY: unsure
 ; LOCATION: (586)
 ; OTHER INFORMATION: n=A, T, C or G
 US-09-796-692-8273

Query Match
 Best Local Similarity 99.7%; Pred. No. 2.1e-141;
 Matches 599; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy	2059	CTGGAAGCCACGAGATATTTGAACTGCTGGGGCGTGACACTGTGACAGCAATTT	2118
Db	1	CTGGAAGCCACGAGATATTTGAACTGCTGGGGCGTGACACTGTGACAGCAATTT	60
Qy	2119	TACTTGATTTTGAATCTGTTGCTATGCTGATCTTCAACTATCTAAGAGTAAAGA	2178
Db	61	TACTTGATTTTGAATCTGTTGCTATGCTGATCTTCAACTATCTAAGAGTAAAGA	120
Qy	2179	GAAAAATTTTCAGGACTTTGGACAGAGATTTTCAAGGAACACAATTTTACCC	2238
Db	121	GAAAAATTTTCAGGACTTTGGACAGAGATTTTCAAGGAACACAATTTTACCC	180
Qy	2239	ACTTCCAAATCACATCCAAATTCAGCATGCTGTTCAAGAGAGATTTGACATACCCG	2298
Db	181	ACTTCCAAATCACATCCAAATTCAGCATGCTGTTCAAGAGAGATTTGACATACCCG	240
Qy	2299	GACTCGGATCAAAATCTCAGGGCTTCATGGGAATTCATTTCACTCTGAAGATGAA	2358
Db	241	GACTCGGATCAAAATCTCAGGGCTTCATGGGAATTCATTTCACTCTGAAGATGAA	300
Qy	2359	TATGAAAAACCAAAAAGGCTTGAAGAGAGAGAGAGACTTGAATGCTTACATTTGA	2418
Db	301	TATGAAAAACCAAAAAGGCTTGAAGAGAGAGAGAGACTTGAATGCTTACATTTGA	360
Qy	2419	CTTCTTTGCTTTGCATATCAAGTTGCCAAAGGAATGGAATTTCTGGAATTTAAGTCGT	2478

Db 361 CTTCTTTGCTTTCATATCAAGTTGCCAAGGAATGGAATTTCTGGAAATTAAGTCGTGT 420
Qy 2479 GTTCACAGACACCTGGCCGCGCAGGAACCTGCTTGTCTACCCACCGGAAAGTGGTGAAGATA 2538
Db 421 GTTCACAGACACCTGGCCGCGCAGGAACCTGCTTGTCTACCCACCGGAAAGTGGTGAAGATA 480
Qy 2539 TGTGACTTTGGATTGGCTCGAGATATCATGAGTGATTTCCAACTATGTTGTCAGGGGCAAT 2598
Db 481 TGTGACTTTGGATTGGCTCGAGATATCATGAGTGATTTCCAACTATGTTGTCAGGGGCAAT 540
Qy 2599 GCCCGTCTGCTGTAAATGATGGCCCGCGAAGACCTGTTGAAGCATCTACACCAAT 2658
Db 541 GCCCGTCTGCTGTAAATGATGGCCCGCGAAGACCTGTTGAAGCATCTACACCAAT 600
Qy 2659 A 2659
Db 601 A 601

RESULT 15
US-10-040-862-8273
; Sequence 8273, Application US/10040862
; Publication No. US20030078396A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; TITLE OF INVENTION: Hematological Malignancies
; FILE REFERENCE: 014058-013520US
; CURRENT APPLICATION NUMBER: US/10/040,862
; CURRENT FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: US 60/223,378
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: US 09/796,692
; PRIOR FILING DATE: 2001-03-01
; NUMBER OF SEQ ID NOS: 10467
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8273
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (586)
; OTHER INFORMATION: n=A,T,C or G
US-10-040-862-8273

Query Match 17.1%; Score 598.4; DB 14; Length 601;
Best Local Similarity 99.7%; Pred. No. 21e-141;
Matches 599; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 2059 CTGGGAAGCCACGAGAAATATTGTGAACCTCTGGGGCGTGCACACTGTGAGGACCAATT 2118
Db 1 CTGGGAAGCCACGAGAAATATTGTGAACCTCTGGGGCGTGCACACTGTGAGGACCAATT 60
Qy 2119 TACTTGATTTTGAATACTGTTGCTATGTTGATCTTCTCAACTATCTTAAGAAAGTAAAGA 2178
Db 61 TACTTGATTTTGAATACTGTTGCTATGTTGATCTTCTCAACTATCTTAAGAAAGTAAAGA 120
Qy 2179 GAAAAATTTTCAACAGGACTTGCACAGAGATTTTCAAGGAAACACAAATTTTCAGTTTTTACCCC 2238
Db 121 GAAAAATTTTCAACAGGACTTGCACAGAGATTTTCAAGGAAACACAAATTTTCAGTTTTTACCCC 180
Qy 2239 ACTTTCCAATCACATCCAAATTTCCAGCATGCTGTGTTCAAGAGAAAGTTCAGATACACCG 2298
Db 181 ACTTTCCAATCACATCCAAATTTCCAGCATGCTGTGTTCAAGAGAAAGTTCAGATACACCG 240
Qy 2299 GACTCGGATCAAAATCTCAGGCTTTCATGGGAATTCATTTCACTCTGAAGATGAATTTGAA 2358
Db 241 GACTCGGATCAAAATCTCAGGCTTTCATGGGAATTCATTTCACTCTGAAGATGAATTTGAA 300
Qy 2359 TATGAAAACCAAAAAGGCTGGGAAGAGAGAGGACTTTGAATGTGCTTACATTTGAAGAT 2418
Db 301 TATGAAAACCAAAAAGGCTGGGAAGAGAGAGGACTTTGAATGTGCTTACATTTGAAGAT 360
Qy 2419 CTTCTTTGCTTTGCATATCAAGTTGCCAAAGGAATGGAATTTCTGGAATTTTAAGTCGTGT 2478
Db 361 CTTCTTTGCTTTGCATATCAAGTTGCCAAAGGAATGGAATTTCTGGAATTTTAAGTCGTGT 420
Qy 2479 GTTCACAGACACCTGGCCGCGCAGGAACGTGCTTGTCTACCCACGGGAAAGTGTGAAGATA 2538
Db 421 GTTCACAGACACCTGGCCGCGCAGGAACGTGCTTGTCTACCCACGGGAAAGTGTGAAGATA 480
Qy 2539 TGTGACTTTGGATTGGCTCGAGATATCATGAGTGATTTCCAACTATGTTGTCAGGGGCAAT 2598
Db 481 TGTGACTTTGGATTGGCTCGAGATATCATGAGTGATTTCCAACTATGTTGTCAGGGGCAAT 540
Qy 2599 GCCCGTCTGCTGTAAATGATGGCCCGCGAAGACCTGTTGAAGCATCTACACCAAT 2658
Db 541 GCCCGTCTGCTGTAAATGATGGCCCGCGAAGACCTGTTGAAGCATCTACACCAAT 600
Qy 2659 A 2659
Db 601 A 601

Search completed: August 28, 2003, 13:26:40
Job time : 1027.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 27, 2003, 15:21:07 ; Search time 284.45 Seconds

(without alignments)
5432.528 Million cell updates/sec

Title: US-09-919-408A-3

Perfect score: 3501

Sequence: 1 CGAGCGCGCATCCGAGGCT.....AAAAAAAAAAAAAAAAAAAA 3501

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued Patents NA:*
- 1: /cgn2_6/ptodata/2/ina/5A COMB.seq:*
 - 2: /cgn2_6/ptodata/2/ina/5B COMB.seq:*
 - 3: /cgn2_6/ptodata/2/ina/6A COMB.seq:*
 - 4: /cgn2_6/ptodata/2/ina/6B COMB.seq:*
 - 5: /cgn2_6/ptodata/2/ina/PCTUS COMB.seq:*
 - 6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3501	100.0	3501	1	US-07-977-451-3
2	3501	100.0	3501	1	US-08-252-517-3
3	3501	100.0	3501	1	US-07-906-397A-3
4	3501	100.0	3501	1	US-08-601-891-3
5	3501	100.0	3501	2	US-09-021-324-3
6	3501	100.0	3501	5	PCT-US92-05401-3
7	3501	100.0	3501	5	PCT-US92-09893-3
8	3475	99.3	3475	1	US-08-222-299-3
9	3475	99.3	3475	2	US-08-434-878-3
10	3475	99.3	3475	5	PCT-US95-03718-3
11	3474.4	99.2	3476	1	US-08-183-211-1
12	3474.4	99.2	3476	5	PCT-US95-00176A-1
13	3083.2	88.1	3120	4	US-08-222-616-22
14	3083.2	88.1	3120	4	US-08-446-648-22
15	3083.2	88.1	3120	5	PCT-US95-04228-22
16	2274.6	65.0	3521	1	US-08-222-299-1
17	2274.6	65.0	3521	2	US-08-434-878-1
18	2274.6	65.0	3521	5	PCT-US95-03718-1
19	2274.6	65.0	3521	5	PCT-US95-03718-1
20	2272.2	63.6	3453	1	US-07-813-593-1
21	2272.2	63.6	3453	1	US-07-977-451-1
22	2272.2	63.6	3453	1	US-07-946-507-1
23	2272.2	63.6	3453	1	US-08-252-517-1
24	2272.2	63.6	3453	1	US-07-906-397A-1
25	2272.2	63.6	3453	1	US-08-601-891-1
26	2272.2	63.6	3453	2	US-09-021-324-1
27	2272.2	63.6	3453	5	PCT-US92-02750-1
28	2272.2	63.6	3453	5	PCT-US92-05401-1

28	2227.2	63.6	3453	5	PCT-US92-09893-1	Sequence 1, Appli
29	624.4	17.8	1894	3	US-07-912-122-3	Sequence 3, Appli
30	624.4	17.8	1894	5	PCT-US93-06404-3	Sequence 3, Appli
31	332	9.5	332	5	PCT-US92-02750-3	Sequence 3, Appli
32	313.2	8.9	5084	1	US-08-306-691B-21	Sequence 21, Appli
33	313.2	8.9	5084	5	PCT-US93-06251-25	Sequence 25, Appli
34	302.6	8.6	4054	1	US-08-180-195-35	Sequence 35, Appli
35	302.6	8.6	4054	1	US-08-477-329-35	Sequence 35, Appli
36	302.6	8.6	4054	2	US-08-475-458-35	Sequence 35, Appli
37	302.6	8.6	4054	3	US-08-980-400-35	Sequence 35, Appli
38	302.6	8.6	4054	3	US-09-583-459A-35	Sequence 35, Appli
39	302.6	8.6	4054	3	US-09-583-210-35	Sequence 35, Appli
40	302.6	8.6	4054	4	US-09-583-449A-35	Sequence 35, Appli
41	302.6	8.6	4054	4	US-09-435-059-35	Sequence 35, Appli
42	302.6	8.6	4100	1	US-08-168-917-3	Sequence 3, Appli
43	302.6	8.6	4100	2	US-08-460-510-3	Sequence 3, Appli
44	302.6	8.6	4100	2	US-08-460-490-3	Sequence 3, Appli
45	302.6	8.6	4100	5	PCT-US92-00730-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1

US-07-977-451-3
; Sequence 3, Application US/07977451
; Patent No. 5270458
; GENERAL INFORMATION:
; APPLICANT: Lemischka, Ihor R.
; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL
; TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ImClone Systems Incorporated
; STREET: 180 Varick Street
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10014
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/977.451
; FILING DATE: 19921119
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US UNASSIGNED
; FILING DATE: 12-NOV-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/906.397
; FILING DATE: 26-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US PCT/US92/05401
; FILING DATE: 26-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: TW 81102961
; FILING DATE: 15-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US PCT/US92/02750
; FILING DATE: 02-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/813.593
; FILING DATE: 24-DEC-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/793.065
; FILING DATE: 15-NOV-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/728.913
; FILING DATE: 28-JUN-1991
; PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/679,666
 FILING DATE: 02-APR-1991
 ATTORNEY/AGENT INFORMATION:
 NAME: Feit, Irving N.
 REGISTRATION NUMBER: 28,601
 REFERENCE/DOCKET NUMBER: LEM-3-7P
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 212-645-1405
 TELEFAX: 212-645-2054
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 3501 base pairs
 TYPE: NUCLEIC ACID
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 FRAGMENT TYPE: N-terminal
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 58...3039
 FEATURE:
 NAME/KEY: mat_peptide
 LOCATION: 139...3036
 FEATURE:
 NAME/KEY: sig_peptide
 LOCATION: 58...138
 US-07-977-451-3

Query Match 100.0%; Score 3501; DB 1; Length 3501;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 3501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	CGAGCGGCATCCGAGCGTGGCGCGCCCTGGGGGACCGGGCTCGGAGGCCATG	60
DB	1	CGAGCGGCATCCGAGCGTGGCGCGCCCTGGGGGACCGGGCTCGGAGGCCATG	60
QY	61	CGGCGTGGCGGACCGCGGACCGTGGCGTCTGTTGTTTCTGCAATGATA	120
DB	61	CGGCGTGGCGGACCGCGGACCGTGGCGTCTGTTGTTTCTGCAATGATA	120
QY	121	TTTGGGACTATTACAAATCAAGATCTGCCTGTCATCAAGTGTGTTTAAATCAATCAAG	180
DB	121	TTTGGGACTATTACAAATCAAGATCTGCCTGTCATCAAGTGTGTTTAAATCAATCAAG	180
QY	181	AACAAATGATTCAGTGGGGAAGTCAATCATATCCATGTCATGATCAAGTGTGTTTAAATCAATCAAG	240
DB	181	AACAAATGATTCAGTGGGGAAGTCAATCATATCCATGTCATGATCAAGTGTGTTTAAATCAATCAAG	240
QY	241	GACCTCGGCTGCGTTGAGACCCAGAGCTCAGGACAGTGTACGAAGCTGCGCTGTG	300
DB	241	GACCTCGGCTGCGTTGAGACCCAGAGCTCAGGACAGTGTACGAAGCTGCGCTGTG	300
QY	301	GAAGTGGATGTATCTGTTCCATCACACTGCAAGTGTGTCGATGCCCGCCAGGGAACATT	360
DB	301	GAAGTGGATGTATCTGTTCCATCACACTGCAAGTGTGTCGATGCCCGCCAGGGAACATT	360
QY	361	TCCTGTCTCGGCTCTTTAAGACACAGCTCCCTGAAATGGCAGCCACATTTGATTACAA	420
DB	361	TCCTGTCTCGGCTCTTTAAGACACAGCTCCCTGAAATGGCAGCCACATTTGATTACAA	420
QY	421	AACAGAGGAGTGTCTTCATGGTCAATTTGAAATGACAGAAACCCAGCTGAGAAATAC	480
DB	421	AACAGAGGAGTGTCTTCATGGTCAATTTGAAATGACAGAAACCCAGCTGAGAAATAC	480
QY	481	CTACTTTTATTCAGAGTGAAGTACCAATTAACAATATTGTTTACAGTGAATAAGA	540
DB	481	CTACTTTTATTCAGAGTGAAGTACCAATTAACAATATTGTTTACAGTGAATAAGA	540
QY	541	AATACCTCTTTTACATTAAGAAGACCTTACTTTAGAAAAATGAAAAACCCAGGACGCC	600
DB	541	AATACCTCTTTTACATTAAGAAGACCTTACTTTAGAAAAATGAAAAACCCAGGACGCC	600

QY	601	CTGCTCTGCATATCTGAGAGCGTTCCAGAGCCGATCGTGGATCGGTGCTTTGCCATTCA	660
DB	601	CTGCTCTGCATATCTGAGAGCGTTCCAGAGCCGATCGTGGATCGGTGCTTTGCCATTCA	660
QY	661	CAGGGGAAAGCTGTAAAGAAAGTCCAGCTGTGTTTAAAGAGGAGGAAAAAGTCTT	720
DB	661	CAGGGGAAAGCTGTAAAGAAAGTCCAGCTGTGTTTAAAGAGGAGGAAAAAGTCTT	720
QY	721	CATGAATTTATTTGGGACGACATAGGTGCTGTGCCAGAAATGAATCTGGCGCAGGAATGC	780
DB	721	CATGAATTTATTTGGGACGACATAGGTGCTGTGCCAGAAATGAATCTGGCGCAGGAATGC	780
QY	781	ACGAGCTGTTTCACAATAGATCTAAATCAAACTCCTCAGACACATTTGCCACAATTTT	840
DB	781	ACGAGCTGTTTCACAATAGATCTAAATCAAACTCCTCAGACACATTTGCCACAATTTT	840
QY	841	CTTAAAGTAGGGGACCCCTTATGGATAAGGTGCAAGCTGTTTCAATGTAACCATGGATTTC	900
DB	841	CTTAAAGTAGGGGACCCCTTATGGATAAGGTGCAAGCTGTTTCAATGTAACCATGGATTTC	900
QY	901	GGGCTCACCTGGGAATTAGAAAAACAAAGCACTCGAGGAGGGCAACTACTTTGAGATGAGT	960
DB	901	GGGCTCACCTGGGAATTAGAAAAACAAAGCACTCGAGGAGGGCAACTACTTTGAGATGAGT	960
QY	961	ACCTATTCAACAAACAGAACTATGATCGGATCTGTTTCTTTTGTATCATCAGTGGCA	1020
DB	961	ACCTATTCAACAAACAGAACTATGATCGGATCTGTTTCTTTTGTATCATCAGTGGCA	1020
QY	1021	AGAAACGACCGGATCTACACTTGTTCCTTCCAAAGCATCCAGTCAATCAGCTTTG	1080
DB	1021	AGAAACGACCGGATCTACACTTGTTCCTTCCAAAGCATCCAGTCAATCAGCTTTG	1080
QY	1081	GTTACCATCTGAGGAAAGGATTTATAAATGTACCAATTCAGTGAAGATTTGAAAT	1140
DB	1081	GTTACCATCTGAGGAAAGGATTTATAAATGTACCAATTCAGTGAAGATTTGAAAT	1140
QY	1141	GACCAATATGAAGAGTGTGTTTCTGTGAGTTTAAAGCTACCCACAAATCAGATGT	1200
DB	1141	GACCAATATGAAGAGTGTGTTTCTGTGAGTTTAAAGCTACCCACAAATCAGATGT	1200
QY	1201	ACGTGGACCTTCTCGAAATCATTTCTTGTGAGCAAAAGGCTTGTATACCGATAC	1260
DB	1201	ACGTGGACCTTCTCGAAATCATTTCTTGTGAGCAAAAGGCTTGTATACCGATAC	1260
QY	1261	AGCATATCAAGTGTGCAATCATAGACACAGCAGGAGAAATATATATTCATGACAGAA	1320
DB	1261	AGCATATCAAGTGTGCAATCATAGACACAGCAGGAGAAATATATATTCATGACAGAA	1320
QY	1321	AATGATGATGCCCAATTTTACCAAAATGTTTACCGTGAATATTAAGAGGAAACCTCAAGTG	1380
DB	1321	AATGATGATGCCCAATTTTACCAAAATGTTTACCGTGAATATTAAGAGGAAACCTCAAGTG	1380
QY	1381	CTCGCAGAGCATTCGGCAAGTCTGAGGCTCTGTTCTCGGATGGATACCATTTACCATCT	1440
DB	1381	CTCGCAGAGCATTCGGCAAGTCTGAGGCTCTGTTCTCGGATGGATACCATTTACCATCT	1440
QY	1441	TGGAAGCTGGAAGAGTGTTCAGCAAGTCTCCCACTGCACAGAGAGATCAAGAGGA	1500
DB	1441	TGGAAGCTGGAAGAGTGTTCAGCAAGTCTCCCACTGCACAGAGAGATCAAGAGGA	1500
QY	1501	GTCCTGGAATAGAAAGGCTTAAAGAAAAAGTGTGGAAGTGTGTCGAGCAGTACTCTA	1560
DB	1501	GTCCTGGAATAGAAAGGCTTAAAGAAAAAGTGTGGAAGTGTGTCGAGCAGTACTCTA	1560
QY	1561	AACATGAGTGAAGCCATAAAAGGTTCTGTTCAAGTGTGTCGATCAATTTCCCTTGGC	1620
DB	1561	AACATGAGTGAAGCCATAAAAGGTTCTGTTCAAGTGTGTCGATCAATTTCCCTTGGC	1620
QY	1621	ACATCTTGTGAGACGATCTTTTAACTCTCCAGGCCCTTCCCTTTTCAATCCAGCAAC	1680
DB	1621	ACATCTTGTGAGACGATCTTTTAACTCTCCAGGCCCTTCCCTTTTCAATCCAGCAAC	1680


```

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/252,517
FILING DATE: 31-OCT-1994
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/977,451
FILING DATE: 19-NOV-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/906,397
FILING DATE: 26-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US PCT/US92/05401
FILING DATE: 26-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: TW 81102961
FILING DATE: 15-APR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US PCT/US92/02750
FILING DATE: 02-APR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/813,593
FILING DATE: 24-DEC-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/793,065
FILING DATE: 15-NOV-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/728,913
FILING DATE: 28-JUN-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/679,666
FILING DATE: 02-APR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Feit, Irving N.
REGISTRATION NUMBER: 28,601
REFERENCE/DOCKET NUMBER: LEM-3-7P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-645-1405
TELEFAX: 212-645-2054
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 3501 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
FEATURE:
NAME/KEY: CDS
LOCATION: 58..3039
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 139..3036
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 58..138
US-08-252-517-3

Query Match 100.0%; Score 3501; DB 1; Length 3501;
Best Local Similarity 100.0%; Pred.No. 0;
Matches 3501; Conservative 0; Mismatches 0; Indels 0; Gaps 0

Qy 1 CGAGGCGGCATCCGAGGCGTGGCCGCGGCCCTCGGGGACCCCGGGCTCCGAGGCGCATG 60
Db 1 CGAGGCGGCATCCGAGGCGTGGCCGCGGCCCTCGGGGACCCCGGGCTCCGAGGCGCATG 60
Qy 61 CCGGCGTTGGCGCGGACCGCGGGACCGTGGCGTCTCGTTGTTTTTCTGCAATGATA 120
Db 61 CCGGCGTTGGCGCGGACCGCGGGACCGTGGCGTCTCGTTGTTTTTCTGCAATGATA 120
Qy 121 TTTGGGACTATTACAAATCAAGATCTGCGCTGTGATCAAGTGTGTTTAAATCAATCAATAG 180

```

QY	1261	AGCATATCCAAAGTTTGTGCAATCATTAAGCACCGACGAGGAGAATATATATTTCCATGCAGAA	1320
DB	1261	AGCATATCCAAAGTTTGTGCAATCATTAAGCACCGACGAGGAGAATATATTTCCATGCAGAA	1320
QY	1321	AATGATGATGCCCAATTTTACCAAAATGTTTCAAGCTGAATATTAAGAAGGAAAACTCAAGTG	1380
DB	1321	AATGATGATGCCCAATTTTACCAAAATGTTTCAAGCTGAATATTAAGAAGGAAAACTCAAGTG	1380
QY	1381	CTCGCAGAAGCATCGGCAAGTCAGGGCGTCTGTGTTTCTCGGATGGATACCCATTTACCATCT	1440
DB	1381	CTCGCAGAAGCATCGGCAAGTCAGGGCGTCTGTGTTTCTCGGATGGATACCCATTTACCATCT	1440
QY	1441	TGCACCTGGGAAGAAGTGTTTCAGACAAGTCTCCCAAACCTGCACAGAAGAGATCAAGAAGGA	1500
DB	1441	TGCACCTGGGAAGAAGTGTTTCAGACAAGTCTCCCAAACCTGCACAGAAGAGATCAAGAAGGA	1500
QY	1501	GTCTGGAATAGAAGGCTTAAACAGAAAAGTGTGTTGGACAGTGGGTGTGAGCAGTACTCTA	1560
DB	1501	GTCTGGAATAGAAGGCTTAAACAGAAAAGTGTGTTGGACAGTGGGTGTGAGCAGTACTCTA	1560
QY	1561	AACATGAGTGAAGCCATAAAGGGTTCCTGGTCAAGTGTGTGTCATACAAATTCCTCTGGC	1620
DB	1561	AACATGAGTGAAGCCATAAAGGGTTCCTGGTCAAGTGTGTGTCATACAAATTCCTCTGGC	1620
QY	1621	ACATCTGTGACAGCATCTTTTAAACCTCTCCAGGCCCTTCCCTTTTCAATCCAAAGACAAC	1680
DB	1621	ACATCTGTGACAGCATCTTTTAAACCTCTCCAGGCCCTTCCCTTTTCAATCCAAAGACAAC	1680
QY	1681	ATCTCATCTTATGCAACAAATTTGGTGTGTGTCCTCTCTTCAATGTCGTTTTAAACCTGTCTA	1740
DB	1681	ATCTCATCTTATGCAACAAATTTGGTGTGTGTCCTCTCTTCAATGTCGTTTTAAACCTGTCTA	1740
QY	1741	ATTGTCACAAGTACAAAAAGCAATTTAGGTATGAAGCCAGCTACAGATGGTACAGGTG	1800
DB	1741	ATTGTCACAAGTACAAAAAGCAATTTAGGTATGAAGCCAGCTACAGATGGTACAGGTG	1800
QY	1801	ACGGCTCCTCAGATATGATGACTTCTACGTGTGATTTTCAGAGAATATGAATGATGATCTC	1860
DB	1801	ACGGCTCCTCAGATATGATGACTTCTACGTGTGATTTTCAGAGAATATGAATGATGATCTC	1860
QY	1861	AAATGGGAGTTTCCAAGAGAAAAATTTAGAGTTTGGGAAGGTACTAGGATCAGGTGCTTTT	1920
DB	1861	AAATGGGAGTTTCCAAGAGAAAAATTTAGAGTTTGGGAAGGTACTAGGATCAGGTGCTTTT	1920
QY	1921	GGAAAAAGTGATGAACGCAACAGCTTATGGAATTAAGCAAAAACAGAGTCTCAATCCAGGTT	1980
DB	1921	GGAAAAAGTGATGAACGCAACAGCTTATGGAATTAAGCAAAAACAGAGTCTCAATCCAGGTT	1980
QY	1981	GCCGTCAAAATGCTGAAAGAAAAAGCAGACAGCTCTGAAAGAGAGGCACATCTGTACAGAA	2040
DB	1981	GCCGTCAAAATGCTGAAAGAAAAAGCAGACAGCTCTGAAAGAGAGGCACATCTGTACAGAA	2040
QY	2041	CTCAAGATGATGACCCAGCTGGGAAGCCACGAGAAATTTGTGAACCTGTCTGGGGGGGTGC	2100
DB	2041	CTCAAGATGATGACCCAGCTGGGAAGCCACGAGAAATTTGTGAACCTGTCTGGGGGGGTGC	2100
QY	2101	ACACTGTGAGGACCAATTTACTTGATTTTTGGAATFACCTGTTGCTATGTTGATCTTCTCAAC	2160
DB	2101	ACACTGTGAGGACCAATTTACTTGATTTTTGGAATFACCTGTTGCTATGTTGATCTTCTCAAC	2160
QY	2161	TATCTAAGAAGTAAAGAGAAAAATTTTTCAGAGGACTTGGACAGAGATTTTCAGGACAC	2220
DB	2161	TATCTAAGAAGTAAAGAGAAAAATTTTTCAGAGGACTTGGACAGAGATTTTCAGGACAC	2220
QY	2221	AAATTCAGTGTTTTACCCCACTTTTCCAAATTCACATCCAAATTCAGCATGCTGTTTCAAGA	2280
DB	2221	AAATTCAGTGTTTTACCCCACTTTTCCAAATTCACATCCAAATTCAGCATGCTGTTTCAAGA	2280
QY	2281	GAAGTTTCAGATACACCCGGACTCTCGGATCAAAATCTCAGGGCTTTCATGGAAATTCATTTTCA	2340
DB	2281	GNAAGTTTCAGATACACCCGGACTCTCGGATCAAAATCTCAGGGCTTTCATGGAAATTCATTTTCA	2340

Qy	2341	TCTGAAAGATGAAATTTGAAATATGAAAACCAAAAAGCGCTCGAAGAGAGAGAGGACTTTGAAAT	2400
Db	2341	TCTGAAAGATGAAATTTGAAATATGAAAACCAAAAAGCGCTCGAAGAGAGAGAGGACTTTGAAAT	2400
Qy	2401	GTGCTTACATTTTGAAAGATCTTTCTTCCTTTGCGATATCAAGTTGCGCAAAAGGAATGGAAATTT	2460
Db	2401	GTGCTTACATTTTGAAAGATCTTTCTTCCTTTGCGATATCAAGTTGCGCAAAAGGAATGGAAATTT	2460
Qy	2461	CTGGAATTTTAAGTCTGTGTGTTCACAGAGACCTGTGGCCGCGAGAACTGTCTGTCTACCCAC	2520
Db	2461	CTGGAATTTTAAGTCTGTGTGTTCACAGAGACCTGTGGCCGCGAGAACTGTCTGTCTACCCAC	2520
Qy	2521	GGGAAAGTGTGTGAAGATATGTGTGACTTTTGGATTTGGCTCTCGAGATATCATGAGTGAATCCCAAC	2580
Db	2521	GGGAAAGTGTGTGAAGATATGTGTGACTTTTGGATTTGGCTCTCGAGATATCATGAGTGAATCCCAAC	2580
Qy	2581	TATGTTGTCTAGGGGCAATGCCCGCTCTGCGCTGTAAATGGAATGGATGGCCGCCGAGAAAGCGCTGTTT	2640
Db	2581	TATGTTGTCTAGGGGCAATGCCCGCTCTGCGCTGTAAATGGAATGGATGGCCGCCGAGAAAGCGCTGTTT	2640
Qy	2641	GAAGGCATCTACACCAATTAAGAGTGTGTCTGTGTATATGGAATATATCTGTGGGAAATC	2700
Db	2641	GAAGGCATCTACACCAATTAAGAGTGTGTCTGTGTATATGGAATATATCTGTGGGAAATC	2700
Qy	2701	TTCTCACTTGGTGTGAATCTTACCTTGGCAATTCGGTTGATGCTTAACCTCTTCTACAACTG	2760
Db	2701	TTCTCACTTGGTGTGAATCTTACCTTGGCAATTCGGTTGATGCTTAACCTCTTCTACAACTG	2760
Qy	2761	ATTCAAATCGAATTTAAATTTGGAATCAGCCCAATTTTATGCTACAGAAGAAATATACATTATA	2820
Db	2761	ATTCAAATCGAATTTAAATTTGGAATCAGCCCAATTTTATGCTACAGAAGAAATATACATTATA	2820
Qy	2821	ATGCAATCTCTCTGGGCTTTTGAATCTCAAGGAAAGCGGCATCTTTCCTTAATTTGACTTCG	2880
Db	2821	ATGCAATCTCTCTGGGCTTTTGAATCTCAAGGAAAGCGGCATCTTTCCTTAATTTGACTTCG	2880
Qy	2881	TTTTTGAAGATGTCAAGTGTGGCAGATGCGAAGAGAGCGATGTATCAGAATGTGGATGGCCGT	2940
Db	2881	TTTTTGAAGATGTCAAGTGTGGCAGATGCGAAGAGAGCGATGTATCAGAATGTGGATGGCCGT	2940
Qy	2941	GTTTTCGGAATGTCTCTACACCTACCAAAACAGGGCAGCTTTTCAGCAGAGAGATGGATTTG	3000
Db	2941	GTTTTCGGAATGTCTCTACACCTACCAAAACAGGGCAGCTTTTCAGCAGAGAGATGGATTTG	3000
Qy	3001	GGGCTACTCTCTCGCAGGCTCAGGTCTGGAAGATTCGTAGAGGAAACAATTTAGTTTTAAGG	3060
Db	3001	GGGCTACTCTCTCGCAGGCTCAGGTCTGGAAGATTCGTAGAGGAAACAATTTAGTTTTAAGG	3060
Qy	3061	ACTTCATCCCTCCACCTATCCCTAACAGGCTGTAGATTAACAAACAAAGATTAAATTTCAAT	3120
Db	3061	ACTTCATCCCTCCACCTATCCCTAACAGGCTGTAGATTAACAAACAAAGATTAAATTTCAAT	3120
Qy	3121	CACATAAAGAAATCTATTATCACTGCTGCTTACCGAGACTTTCTCTAGAGGCGCTCT	3180
Db	3121	CACATAAAGAAATCTATTATCACTGCTGCTTACCGAGACTTTCTCTAGAGGCGCTCT	3180
Qy	3181	CGGTTTTACTCTGTGTTTTCAAAAGGACATTTTGTAAATCAAAATCATCTGTACAAAGGCAG	3240
Db	3181	CGGTTTTACTCTGTGTTTTCAAAAGGACATTTTGTAAATCAAAATCATCTGTACAAAGGCAG	3240
Qy	3241	GAGGAGCTGATAATGAACTTTTATTTGGAGCAATGTATCTGCATCCAAAGGCGCTTCTCAGGCCG	3300
Db	3241	GAGGAGCTGATAATGAACTTTTATTTGGAGCAATGTATCTGCATCCAAAGGCGCTTCTCAGGCCG	3300
Qy	3301	GCTTGAGTGAATTTGTACCTGAAGTACAGTATATTTCTGTAAATACATAAAACAAAGC	3360
Db	3301	GCTTGAGTGAATTTGTACCTGAAGTACAGTATATTTCTGTAAATACATAAAACAAAGC	3360
Qy	3361	ATTTTGTCTAAGGAGAGCTAATATGATTTTTTAAAGTCTATCTGTTTTTAAATAATATGTAAA	3420
Db	3361	ATTTTGTCTAAGGAGAGCTAATATGATTTTTTAAAGTCTATCTGTTTTTAAATAATATGTAAA	3420
Qy	3421	TTTTTTTCAGCTATTTTAGTGATATATTTTTATGCGGTGGGAATAAAATTTTCTACACAGAAAA	3480

```

Db      3421 TTTTTCAGCTATTAGTGATATATTTATGGTGGGAATAAAATTTCTACTACAGAAAA 3480
Qy      3481 AAAAAAAAAAAAAAAAAAAAAA 3501
Db      3481 AAAAAAAAAAAAAAAAAAAAAA 3501

RESULT 3
US-07-906-397A-3
; Sequence 3, Application US/07906397A
; Patent No. 5621090
; GENERAL INFORMATION:
; APPLICANT: Lemischka, Ihor R.
; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL
; TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCLONE SYSTEMS INCORPORATED
; STREET: 180 VARICK STREET
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: U.S.A.
; ZIP: 10014
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/906,397A
; FILING DATE: 19920626
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/813,593
; FILING DATE: 24-DEC-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/793,065
; FILING DATE: 15-NOV-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/728,913
; FILING DATE: 28-JUN-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/679,666
; FILING DATE: 02-APR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Felt, Irving N.
; REGISTRATION NUMBER: 28,601
; REFERENCE/DOCKET NUMBER: LEM-3-PPPPPP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-645-1405
; TELEFAX: 212-645-2054
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3501 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 58..3039
; US-07-906-397A-3

Query Match      100.0%; Score 3501; DB 1; Length 3501;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 CGAGGCGGCATCCGAGGGCTGGGCGGCGCCCTCGGGGACCCCGGGTTCGAGGCCATG 60
Db      1 CGAGGCGGCATCCGAGGGCTGGGCGGCGCCCTCGGGGACCCCGGGTTCGAGGCCATG 60

```

[illegible]

Db 3301 GCTTGAGTGAATTGTGACCTGAAGTACAGTATATTTCTTGTAATACATAAAACAAAGC 3360
 Qy 3361 ATTTTGTCTAAGAGAGCTAATATGATTTTAACTCTATGTTTAAATATATATGTA 3420
 Db 3361 ATTTTGTCTAAGAGAGCTAATATGATTTTAACTCTATGTTTAAATATATATGTA 3420
 Qy 3421 TTTTTCAGCTATTAGTCATATATTTTATGGTGGGAATAAAATTTCTACTACAGAAA 3480
 Db 3421 TTTTTCAGCTATTAGTCATATATTTTATGGTGGGAATAAAATTTCTACTACAGAAA 3480
 Qy 3481 AAAAAAAAAAAAAAAAAAAAAA 3501
 Db 3481 AAAAAAAAAAAAAAAAAAAAAA 3501

RESULT 4

US-08-601-891-3
 ; Sequence 3, Application US/08601891
 ; Patent No. 5747651
 ; GENERAL INFORMATION:
 ; APPLICANT: Lemischka, Ihor R.
 ; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL
 ; TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS
 ; NUMBER OF SEQUENCES: 10
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: ImClone Systems Incorporated
 ; STREET: 180 Varick Street
 ; CITY: New York
 ; STATE: New York
 ; COUNTRY: U.S.A.
 ; ZIP: 10014
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/601,891
 ; FILING DATE: 15-FEB-1996
 ; CLASSIFICATION: 530
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/977,451
 ; FILING DATE: 19-NOV-1992
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/906,397
 ; FILING DATE: 26-JUN-1992
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US PCT/US92/05401
 ; FILING DATE: 26-JUN-1992
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: TW 81102961
 ; FILING DATE: 15-APR-1992
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US PCT/US92/02750
 ; FILING DATE: 02-APR-1992
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/813,593
 ; FILING DATE: 24-DEC-1991
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/793,665
 ; FILING DATE: 15-NOV-1991
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/728,913
 ; FILING DATE: 28-JUN-1991
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/679,666
 ; FILING DATE: 02-APR-1991
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Feit, Irving N.
 ; REGISTRATION NUMBER: 28,601
 ; REFERENCE/DOCKET NUMBER: LEM-3-7P
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 212-645-1405

TELEFAX: 212-645-2054
 ; INFORMATION FOR SEQ ID NO: 3:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 3501 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA
 ; HYPOTHEetical: NO
 ; ANTI-SENSE: NO
 ; FRAGMENT TYPE: N-terminal
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 58...3039
 ; FEATURE:
 ; NAME/KEY: mat_peptide
 ; LOCATION: 139...3036
 ; FEATURE:
 ; NAME/KEY: sig_peptide
 ; LOCATION: 58...138
 ; US-08-601-891-3

Query Match 100.0%; Score 3501; DB 1; Length 3501;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 3501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGAGCGGCGATCCGAGGCGTGGCGCGCGCCCTGGGGACCCCGGGCTCCGGAGGCCATG 60
 Db 1 CGAGCGGCGATCCGAGGCGTGGCGCGCGCCCTGGGGACCCCGGGCTCCGGAGGCCATG 60
 Qy 61 CGGGGTTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 120
 Db 61 CGGGGTTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 120
 Qy 121 TTTGGGACTATTACAAATCAAGATCTGCTGCTGATCAAGTGTGTTTAAATCAATCAAG 180
 Db 121 TTTGGGACTATTACAAATCAAGATCTGCTGCTGATCAAGTGTGTTTAAATCAATCAAG 180
 Qy 181 AACAAATGATTCATCAGTGGGGAAGTCAATCATATATCCATGGTATCAGAAATCCCGGAA 240
 Db 181 AACAAATGATTCATCAGTGGGGAAGTCAATCATATATCCATGGTATCAGAAATCCCGGAA 240
 Qy 241 GACCTCGGGTGTGCTTGAGACCCCGAGAGCTCAGGACAGTGTACGAAGTCCCGCTGTG 300
 Db 241 GACCTCGGGTGTGCTTGAGACCCCGAGAGCTCAGGACAGTGTACGAAGTCCCGCTGTG 300
 Qy 301 GAAGTGAATGTATCTGCTTCCATCACTGCAAGTGTGCTGATGCCCGGGAACATT 360
 Db 301 GAAGTGAATGTATCTGCTTCCATCACTGCAAGTGTGCTGATGCCCGGGAACATT 360
 Qy 361 TCCTGTCTCTGGGCTTTTAAAGCAGAGCTCCCTGAATTCAGGACCAATTTTGAATTTACAA 420
 Db 361 TCCTGTCTCTGGGCTTTTAAAGCAGAGCTCCCTGAATTCAGGACCAATTTTGAATTTACAA 420
 Qy 421 AACAGAGAGTGTGTTTCCATGGTCAATTTTGAATAATGACAGAAACCCAGAGTGGAGATAC 480
 Db 421 AACAGAGAGTGTGTTTCCATGGTCAATTTTGAATAATGACAGAAACCCAGAGTGGAGATAC 480
 Qy 481 CTACTTTTATTTCAGAGTGAAGTACCAATTTACAAATTTTGTACAGTGAATATAGA 540
 Db 481 CTACTTTTATTTCAGAGTGAAGTACCAATTTACAAATTTTGTACAGTGAATATAGA 540
 Qy 541 AATACCCCTGCTTTACACATTAAGAGAGCTTACTTTAGAAAAATGAAAAACAGGACGCC 600
 Db 541 AATACCCCTGCTTTACACATTAAGAGAGCTTACTTTAGAAAAATGAAAAACAGGACGCC 600
 Qy 601 CTGGTCTGCATATCTGAGAGCGTTCCAGAGCCGATCGTGGAAATGGGTCTTTGGCATTC 660
 Db 601 CTGGTCTGCATATCTGAGAGCGTTCCAGAGCCGATCGTGGAAATGGGTCTTTGGCATTC 660
 Qy 661 CAGGGGGAAGCTGTAAAGAGAAAGTCCAGCTGTGTTTAAAGAGGAGGAAAGTCTT 720
 Db 661 CAGGGGGAAGCTGTAAAGAGAAAGTCCAGCTGTGTTTAAAGAGGAGGAAAGTCTT 720

Db 241 GACCTCGGCTGTGCTTTGAGACCCAGAGCTCAGGACAGTGTACGAAGCTGCGCTGTG 300
Qy 301 GAATGATCTATCTCTTCATCACACTGCAAGTGTCTGTCGATCCCGAGGAAATTT 360
Db 301 GAATGATCTATCTCTTCATCACACTGCAAGTGTCTGTCGATCCCGAGGAAATTT 360
Qy 361 TCCTGTCTCTGGGTCTTTAAAGCAGAGCTCCCTGAAATGCCAGCCACATTTGATTACAA 420
Db 361 TCCTGTCTCTGGGTCTTTAAAGCAGAGCTCCCTGAAATGCCAGCCACATTTGATTACAA 420
Qy 421 AACAGAGGAGTGTTCCTCATCTGTTTGAATATGACAGAAACCCAGCTGGAGATAC 480
Db 421 AACAGAGGAGTGTTCCTCATCTGTTTGAATATGACAGAAACCCAGCTGGAGATAC 480
Qy 481 CTACTTTTATTCAGAGTGAAGCTTACCAATTTACACAATTTGTTTACAGTGAATAGA 540
Db 481 CTACTTTTATTCAGAGTGAAGCTTACCAATTTACACAATTTGTTTACAGTGAATAGA 540
Qy 541 AATACCCTGCTTTACATTAAGAAGACCTTACTTTAGAAAATGAAAAACCCAGAGGCC 600
Db 541 AATACCCTGCTTTACATTAAGAAGACCTTACTTTAGAAAATGAAAAACCCAGAGGCC 600
Qy 601 CTGGTGTGCATATCTGAGAGGTTCCAGAGCCGATCGTGAATGGGTGCTTTGCGATTCA 660
Db 601 CTGGTGTGCATATCTGAGAGGTTCCAGAGCCGATCGTGAATGGGTGCTTTGCGATTCA 660
Qy 661 CAGGGGAAAGCTGTAAGAAGAAAGTCCAGCTGTTTAAAGAGGAAAGTCTT 720
Db 661 CAGGGGAAAGCTGTAAGAAGAAAGTCCAGCTGTTTAAAGAGGAAAGTCTT 720
Qy 721 CATGAATTTTGGGACGGACATAGGTGTGTCAGAGAAATGAATGAGGAGGAAATGC 780
Db 721 CATGAATTTTGGGACGGACATAGGTGTGTCAGAGAAATGAATGAGGAGGAAATGC 780
Qy 781 ACCAGGCTGTTCAATPAGATCTAAATCAAACTCCTCAGACCAATTTGCCAATTTATTT 840
Db 781 ACCAGGCTGTTCAATPAGATCTAAATCAAACTCCTCAGACCAATTTGCCAATTTATTT 840
Qy 841 CTTAAAGTGGGAAACCTTATGATGAAGTGAAGCTGTTCAATGAGAACCATGATTC 900
Db 841 CTTAAAGTGGGAAACCTTATGATGAAGTGAAGCTGTTCAATGAGAACCATGATTC 900
Qy 901 GGCGTCACTGGGAATTTAGAAAACAAAGCACTCCAGAGGCGCACTACTTTGAGATGAGT 960
Db 901 GGCGTCACTGGGAATTTAGAAAACAAAGCACTCCAGAGGCGCACTACTTTGAGATGAGT 960
Qy 961 ACCTATTCAACAAACAGAACTATGATACGGATTCCTGTTGCTTTGATCATCAGTGGCA 1020
Db 961 ACCTATTCAACAAACAGAACTATGATACGGATTCCTGTTGCTTTGATCATCAGTGGCA 1020
Qy 1021 AGAAACGACACCGGATACACTGTTGCTCTTTCAAGCAATCCAGTCAATCAGCTTTG 1080
Db 1021 AGAAACGACACCGGATACACTGTTGCTCTTTCAAGCAATCCAGTCAATCAGCTTTG 1080
Qy 1081 GTTACCATCTGAGAAAGGATTTATAATGCTACCAATTCAGTGAAGATTATGAATTT 1140
Db 1081 GTTACCATCTGAGAAAGGATTTATAATGCTACCAATTCAGTGAAGATTATGAATTT 1140
Qy 1141 GACCAATATCAAGAGTTTGTGTTTCTGTAGGTTTAAAGCTTACCACAAATCAGATGT 1200
Db 1141 GACCAATATCAAGAGTTTGTGTTTCTGTAGGTTTAAAGCTTACCACAAATCAGATGT 1200
Qy 1201 ACCTGGACCTTCTCGAAATCAATTTCTTTGAGCAAAAGGCTTTGATTAACGGATAC 1260
Db 1201 ACCTGGACCTTCTCGAAATCAATTTCTTTGAGCAAAAGGCTTTGATTAACGGATAC 1260
Qy 1261 AGCATATCCAAAGTTTGCATATCAAGCAACAGCCAGGAGATATATATTCATGCGAAA 1320
Db 1261 AGCATATCCAAAGTTTGCATATCAAGCAACAGCCAGGAGATATATATTCATGCGAAA 1320
Qy 1321 AATGATGATCCCAATTTACCAAAATGTTTCAAGCTGAATATAAGAGGAAACCTCAAGTG 1380
Db 1321 AATGATGATCCCAATTTACCAAAATGTTTCAAGCTGAATATAAGAGGAAACCTCAAGTG 1380

Qy 1381 CTGCGAAGCATCGCAAGTCAGGCGTCTCTGTTCTCGATGATACCCATTACCATCT 1440
Db 1381 CTGCGAAGCATCGCAAGTCAGGCGTCTCTGTTCTCGATGATACCCATTACCATCT 1440
Qy 1441 TGACCTTGAAGAGTGTTCAGACAGTCTCCCAACTGCAAGAGAGATCACAGAGGA 1500
Db 1441 TGACCTTGAAGAGTGTTCAGACAGTCTCCCAACTGCAAGAGAGATCACAGAGGA 1500
Qy 1501 GTCTGGAATAGAAGGCTTACAGAAAGTGTGGAAGTGGGTGTCGAGCAGTACTCTA 1560
Db 1501 GTCTGGAATAGAAGGCTTACAGAAAGTGTGGAAGTGGGTGTCGAGCAGTACTCTA 1560
Qy 1561 AACATGAGTCAAGCCATAAAAGGCTTCTGCTCAAGTGTGTCATACAAATTCCTTGGC 1620
Db 1561 AACATGAGTCAAGCCATAAAAGGCTTCTGCTCAAGTGTGTCATACAAATTCCTTGGC 1620
Qy 1621 ACATCTTCTGAGACGATCCCTTTAAACTCTCCAGGCCCTTCCCTTCAACAGACAAC 1680
Db 1621 ACATCTTCTGAGACGATCCCTTTAAACTCTCCAGGCCCTTCCCTTCAACAGACAAC 1680
Qy 1681 ATCTCATTTCTATGCAAAATTTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1740
Db 1681 ATCTCATTTCTATGCAAAATTTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1740
Qy 1741 ATTTGTCAAGTACAAAAGCAATTTAGTATGAAAGCCAGCTACAGATGATGATGATG 1800
Db 1741 ATTTGTCAAGTACAAAAGCAATTTAGTATGAAAGCCAGCTACAGATGATGATGATG 1800
Qy 1801 ACCGGCTCTCAGATAATGATGATCTCTACGTTGATTTTCAAGAGATATGATGATCTC 1860
Db 1801 ACCGGCTCTCAGATAATGATGATCTCTACGTTGATTTTCAAGAGATATGATGATCTC 1860
Qy 1861 AATGGAGTTTCAAGAGAAATTTAGAGTTTGGGAAAGTACTAGAGTCAAGTGTCTTTT 1920
Db 1861 AATGGAGTTTCAAGAGAAATTTAGAGTTTGGGAAAGTACTAGAGTCAAGTGTCTTTT 1920
Qy 1921 GGAAGTGTATGACGCAACAGCTTATGGAATTTAGCAAAACAGAGTCTCAATCCAGGT 1980
Db 1921 GGAAGTGTATGACGCAACAGCTTATGGAATTTAGCAAAACAGAGTCTCAATCCAGGT 1980
Qy 1981 GCGTCAAAATGCTGAAAGAAAGCAGACAGCTCTGAAAGAGAGGCACTCATGTCAAG 2040
Db 1981 GCGTCAAAATGCTGAAAGAAAGCAGACAGCTCTGAAAGAGAGGCACTCATGTCAAG 2040
Qy 2041 CTCAGATGATGACCCAGCTGGGAAAGCCAGAGATATTTGAAACCTGCTGGGGCGTGC 2100
Db 2041 CTCAGATGATGACCCAGCTGGGAAAGCCAGAGATATTTGAAACCTGCTGGGGCGTGC 2100
Qy 2101 ACATGTCAAGCAAAATTTACTGATTTTGAATACTGTTGCTATGTTGATCTTCAAC 2160
Db 2101 ACATGTCAAGCAAAATTTACTGATTTTGAATACTGTTGCTATGTTGATCTTCAAC 2160
Qy 2161 TATCTAAGAGTAAAGAGAAATTTTCAAGAGCTTTGGAAGAGATTTTCAAGAGAAC 2220
Db 2161 TATCTAAGAGTAAAGAGAAATTTTCAAGAGCTTTGGAAGAGATTTTCAAGAGAAC 2220
Qy 2221 AATTTCAAGTTTACCCCACTTTCCAAATCACATCCAAATTTCCAGATGCTGTTCAAGA 2280
Db 2221 AATTTCAAGTTTACCCCACTTTCCAAATCACATCCAAATTTCCAGATGCTGTTCAAGA 2280
Qy 2281 GAAGTTCAGATACACCGGACTCGGATCAAAATCTCAGGGCTTTCATGGGAATTTCAAC 2340
Db 2281 GAAGTTCAGATACACCGGACTCGGATCAAAATCTCAGGGCTTTCATGGGAATTTCAAC 2340
Qy 2341 TCTGAAGATGAAATTTGAATATGAAACCAAAAGGCTGGAAGAGGAGGATTTGAAT 2400
Db 2341 TCTGAAGATGAAATTTGAATATGAAACCAAAAGGCTGGAAGAGGAGGATTTGAAT 2400
Qy 2401 GTGCTTACATTTGAGATCTTTTGTGCTATCATCAAGTTGCAAGAGGATGGAATTT 2460
Db 2401 GTGCTTACATTTGAGATCTTTTGTGCTATCATCAAGTTGCAAGAGGATGGAATTT 2460

Qy	2461	CTGGAATTTAAGTCGTGTGTTCTACAGAGACCTGGCCGCCACAGAACGTCGCTTGTCACCCAC	2520
Db	2461	CTGGAATTTAAGTCGTGTGTTCTACAGAGACCTGGCCGCCACAGAACGTCGCTTGTCACCCAC	2520
Qy	2521	GGGAAAGTGGTGAAGATATGTGATCTTTGGATTTGGCTCGAGATATCATGTAGTGATTTCCAAC	2580
Db	2521	GGGAAAGTGGTGAAGATATGTGATCTTTGGATTTGGCTCGAGATATCATGTAGTGATTTCCAAC	2580
Qy	2581	TATGTTGTCTAGGGGCAATGCCGCTCCCTGTGTAATTAAGATGGCCCCCGGAAAGCCTGTTTT	2640
Db	2581	TATGTTGTCTAGGGGCAATGCCGCTCCCTGTGTAATTAAGATGGCCCCCGGAAAGCCTGTTTT	2640
Qy	2641	GAAGGCATCTACACCAATTAAGAGTGATGTCTGGTTCATATGGAATATTACTGTGGGAAATC	2700
Db	2641	GAAGGCATCTACACCAATTAAGAGTGATGTCTGGTTCATATGGAATATTACTGTGGGAAATC	2700
Qy	2701	TTCTCACTTTGGTGTGAATCCTTACCCTGGCAATCCGGTTGATGCTAACTTCTTCAAACTG	2760
Db	2701	TTCTCACTTTGGTGTGAATCCTTACCCTGGCAATCCGGTTGATGCTAACTTCTTCAAACTG	2760
Qy	2761	ATTCAAAAATGGATTTAAAAATGGATCAGCCATTTTATGCTACAGAAAGAAATATACATTATA	2820
Db	2761	ATTCAAAAATGGATTTAAAAATGGATCAGCCATTTTATGCTACAGAAAGAAATATACATTATA	2820
Qy	2821	ATGCAATCTCTCTGGGGCTTTTGACTCAAGGAAACGGCCATCCTTCCCTAAATTTGACTTCG	2880
Db	2821	ATGCAATCTCTCTGGGGCTTTTGACTCAAGGAAACGGCCATCCTTCCCTAAATTTGACTTCG	2880
Qy	2881	TTTTTATGAGATGTCTAGCTGGCAGATGCGAGAGACGGATGTATCAGATGTGGATGGCCGT	2940
Db	2881	TTTTTATGAGATGTCTAGCTGGCAGATGCGAGAGACGGATGTATCAGATGTGGATGGCCGT	2940
Qy	2941	GTTTCCGAATGTCTCTCACCTCACCTCAAAAACAGGGCGACCTTTTCAGCAGAGAGATGGATTTG	3000
Db	2941	GTTTCCGAATGTCTCTCACCTCACCTCAAAAACAGGGCGACCTTTTCAGCAGAGAGATGGATTTG	3000
Qy	3001	GGGCTACTCTCTCCGCGAGGCTCAGGTCCGAGATTCGTAGAGGAAACAATTTAGTTTTAAGG	3060
Db	3001	GGGCTACTCTCTCCGCGAGGCTCAGGTCCGAGATTCGTAGAGGAAACAATTTAGTTTTAAGG	3060
Qy	3061	ACTTCATCCCTCCACCTATCCCTAACAGGCTGTAGATTACCAAAACAAGATTTAATTTTCAT	3120
Db	3061	ACTTCATCCCTCCACCTATCCCTAACAGGCTGTAGATTACCAAAACAAGATTTAATTTTCAT	3120
Qy	3121	CACATAAAGAAATCTATTATCAACTCTGCTTCCACGAGCTTTCTCTAGAGCCGCTCT	3180
Db	3121	CACATAAAGAAATCTATTATCAACTCTGCTTCCACGAGCTTTCTCTAGAGCCGCTCT	3180
Qy	3181	CGGTTTACTCTGTGTTTTCAAAAGGACCTTTTGTGAAATCAAAATCATCTGTCCACAGGCGAG	3240
Db	3181	CGGTTTACTCTGTGTTTTCAAAAGGACCTTTTGTGAAATCAAAATCATCTGTCCACAGGCGAG	3240
Qy	3241	GAGGAGCTGATAATGAACCTTTATTGGAGCATTTGATCTGCATCCAGGCTTCTCAGGCCG	3300
Db	3241	GAGGAGCTGATAATGAACCTTTATTGGAGCATTTGATCTGCATCCAGGCTTCTCAGGCCG	3300
Qy	3301	GCTTGAGTGAAATGTGTACCTGAAGTACAGTATATTCTGTGAAATACATAAACAACAAAGC	3360
Db	3301	GCTTGAGTGAAATGTGTACCTGAAGTACAGTATATTCTGTGAAATACATAAACAACAAAGC	3360
Qy	3361	ATTTTGTCTAAGGAGAAGCTAATATGATTTTTTTAAGTCTATGCTTTTTAAAAATAATATGTAAA	3420
Db	3361	ATTTTGTCTAAGGAGAAGCTAATATGATTTTTTTAAGTCTATGCTTTTTAAAAATAATATGTAAA	3420
Qy	3421	TTTTTTCAGCTATTTAGTGATATATTTTTATGGGTGGGAATAAAATTTCTACTACAGAAAAA	3480
Db	3421	TTTTTTCAGCTATTTAGTGATATATTTTTATGGGTGGGAATAAAATTTCTACTACAGAAAAA	3480
Qy	3481	AAAAAAAAAAAAAAAAAAAAA 3501	
Db	3481	AAAAAAAAAAAAAAAAAAAAA 3501	

RESULT 6
PCT-US92-05401-3
Sequence 3, Application PC/TUS9205401
GENERAL INFORMATION:
APPLICANT: Lemischka, Ihor R.
TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL
TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: IMCLONE SYSTEMS INCORPORATED
CITY: 180 VARICK STREET
STATE: NEW YORK
COUNTRY: U.S.A.
ZIP: 10014
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Review #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/05401
FILING DATE: 19920626
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Feit, Irving N.
REGISTRATION NUMBER: 28,601
REFERENCE/DOCKET NUMBER: LEM-3-PPPPPT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-645-1405
TELEFAX: 212-645-2054
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 3501 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cdna
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: join(1..3036, 3040..3111, 3115..3123, 3127..
LOCATION: 3172..3303, 3307..3348, 3352..3378, 3382..
PCT-US92-05401-3

	Query Match	100.0.0%;	Score 3501;	DB 5;	Length 3501;
	Best Local Similarity	100.0.0%;	Pred. No. 0;		
	Matches 3501;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	CGAGGCGGCATCGAGGGCTGGGCGGCGCCCTGGGGGACCCCGGGCTCCGGAGGCCATG	60		
Db	1	CGAGGCGGCATCGAGGGCTGGGCGGCGCCCTGGGGGACCCCGGGCTCCGGAGGCCATG	60		
Qy	61	CCGCGCTTGGCGGGCGACGCGGGGACCGTGCCGCTGCTCGTGTGTTTTCTGCAATGATA	120		
Db	61	CCGCGCTTGGCGGGCGAGCGGGGACCGTGCCGCTGCTCGTGTGTTTTCTGCAATGATA	120		
Qy	121	TTTGGGACTATTCAAATCAAAGATCTGCCTGTGTGATCAAGTGTGTTTAAATCAATCATAAG	180		
Db	121	TTTGGGACTATTCAAATCAAAGATCTGCCTGTGTGATCAAGTGTGTTTAAATCAATCATAAG	180		
Qy	181	AACAAATGATTCATCAGTGGGGAGTCAATCATATCCCATGTGTATCAGAAATCCCGGAA	240		
Db	181	AACAAATGATTCATCAGTGGGGAGTCAATCATATCCCATGTGTATCAGAAATCCCGGAA	240		
Qy	241	GACCTCGGGTGTGCGTTTGAGACCCGAGCTCAGGGACAGTGTACGAAAGTGCCTGCTGTG	300		
Db	241	GACCTCGGGTGTGCGTTTGAGACCCGAGCTCAGGGACAGTGTACGAAAGTGCCTGCTGTG	300		
Qy	301	GAAGTGATGTATCTGCTTTCATACACACTGCAAGTGTGCTCGATGCCCCAGGGAAACATT	360		
Db	301	GAAGTGATGTATCTGCTTTCATACACACTGCAAGTGTGCTCGATGCCCCAGGGAAACATT	360		

QY 361 TCCTGTCTCTGGGCTTTTAAGCAGACAGCTCCCTGAATTCGCCAGCCACATTTTGGATTACAA 420
DB |||||
DB 361 TCCTGTCTCTGGGCTTTTAAGCAGACAGCTCCCTGAATTCGCCAGCCACATTTTGGATTACAA 420
QY 421 AACAGAGAGGTGTTTCCATGGTCAATTTTGAATAATGACAGAAACCCAAAGCTGGAGAAATAC 480
DB |||||
DB 421 AACAGAGAGGTGTTTCCATGGTCAATTTTGAATAATGACAGAAACCCAAAGCTGGAGAAATAC 480
QY 481 CTACTTTTATTACAGAGTGAAGCTACCAATTTACACAATATTGTTTACAGTGAGTATAGA 540
DB |||||
DB 481 CTACTTTTATTACAGAGTGAAGCTACCAATTTACACAATATTGTTTACAGTGAGTATAGA 540
QY 541 AATACCCCTGCTTTACACATTAAGAAAGACCTTTACTTTTGAATAATGGAATAACAGAGGCC 600
DB |||||
DB 541 AATACCCCTGCTTTACACATTAAGAAAGACCTTTACTTTTGAATAATGGAATAACAGAGGCC 600
QY 601 CTGCTCTGCAATCTGAGAGCGTTCCAGAGCGATCGTGGAAATGGTCTTTTGGCAATTC 660
DB |||||
DB 601 CTGCTCTGCAATCTGAGAGCGTTCCAGAGCGATCGTGGAAATGGTCTTTTGGCAATTC 660
QY 661 CAGGGGAAAGCTGTAAGAAAGAAAGTCCAGCTGTTGTTTAAAGAGGAGGAAAGTGCCT 720
DB |||||
DB 661 CAGGGGAAAGCTGTAAGAAAGAAAGTCCAGCTGTTGTTTAAAGAGGAGGAAAGTGCCT 720
QY 721 CATGAATTTATTGGGACGGACATAGGTGCTGTCAGAAATGAATGAGTGGGAGGAAATGC 780
DB |||||
DB 721 CATGAATTTATTGGGACGGACATAGGTGCTGTCAGAAATGAATGAGTGGGAGGAAATGC 780
QY 781 ACCAGGCTGTTTCACAATAGATCTAAATCAAACTCTCAGACCAATTCGCCACAATTTAT 840
DB |||||
DB 781 ACCAGGCTGTTTCACAATAGATCTAAATCAAACTCTCAGACCAATTCGCCACAATTTAT 840
QY 841 CTTAAAGTAGGGAAACCTTTATGATAGGTGCAAGCTGTTTCAATGTAACCAATGGAATTC 900
DB |||||
DB 841 CTTAAAGTAGGGAAACCTTTATGATAGGTGCAAGCTGTTTCAATGTAACCAATGGAATTC 900
QY 901 GGGCTACCTGGGAAATTTAGAAACAAAGCACTCAGAGGCGCACTACTTTGAGATGAGT 960
DB |||||
DB 901 GGGCTACCTGGGAAATTTAGAAACAAAGCACTCAGAGGCGCACTACTTTGAGATGAGT 960
QY 961 ACCTATTCAACAAACAGAACTATGATACGGATTCCTGTTTGTGTTTGTATCATCAGTGGCA 1020
DB |||||
DB 961 ACCTATTCAACAAACAGAACTATGATACGGATTCCTGTTTGTGTTTGTATCATCAGTGGCA 1020
QY 1021 AGAAACGACACCGGATACACTTGTTCCTTTCAAGCATCCAGTCAATCAGCTTTG 1080
DB |||||
DB 1021 AGAAACGACACCGGATACACTTGTTCCTTTCAAGCATCCAGTCAATCAGCTTTG 1080
QY 1081 GTTACCATCGTAGAAAGGGAATTTATAATGCTACCAATTTCAAGTGAAGATTTGAAAT 1140
DB |||||
DB 1081 GTTACCATCGTAGAAAGGGAATTTATAATGCTACCAATTTCAAGTGAAGATTTGAAAT 1140
QY 1141 GACCAATATGAAGATTTTGTGTTTCTGTCAGGTTTAAAGCTTACCCACAATCAGATGT 1200
DB |||||
DB 1141 GACCAATATGAAGATTTTGTGTTTCTGTCAGGTTTAAAGCTTACCCACAATCAGATGT 1200
QY 1201 ACSTGGACCTTCTCTCGAAATCAATTTCTTGTGAGCAAAAGGCTTTGATPAAACGGATAC 1260
DB |||||
DB 1201 ACSTGGACCTTCTCTCGAAATCAATTTCTTGTGAGCAAAAGGCTTTGATPAAACGGATAC 1260
QY 1261 AGCATATCCAAGTTTTCGAAATCATTAAGCAGCAGCCAGGAGAAATATATATTCATGCGAAA 1320
DB |||||
DB 1261 AGCATATCCAAGTTTTCGAAATCATTAAGCAGCAGCCAGGAGAAATATATATTCATGCGAAA 1320
QY 1321 AATGATGATGCCAAATTTTACCATAATGTTTCAACGCTGAATATAGAAGGAACCTCAAGTG 1380
DB |||||
DB 1321 AATGATGATGCCAAATTTTACCATAATGTTTCAACGCTGAATATAGAAGGAACCTCAAGTG 1380
QY 1381 CTGCGAAGCATCGGCAAGTCAGGCGTCTGTTTCTCGGATGATACCCATTAACCATCT 1440
DB |||||
DB 1381 CTGCGAAGCATCGGCAAGTCAGGCGTCTGTTTCTCGGATGATACCCATTAACCATCT 1440

QY 1441 TGGACCTGGAAGAGTGTTCAGACAAAGTCTCCAACTGCACAGAAAGATCACAGAAGGA 1500
DB |||||
DB 1441 TGGACCTGGAAGAGTGTTCAGACAAAGTCTCCAACTGCACAGAAAGATCACAGAAGGA 1500
QY 1501 GTCTGGAATAGAAAGGCTTAAACAGAAAAGTGTGTGGACAGTGGGTGTGCGAGCAGTACTCTA 1560
DB |||||
DB 1501 GTCTGGAATAGAAAGGCTTAAACAGAAAAGTGTGTGGACAGTGGGTGTGCGAGCAGTACTCTA 1560
QY 1561 AACATGAGTGAAGCCATFAAAGGGTTCCTGGTCAAGTGTGTGCATACAAATTCCTCTGGC 1620
DB |||||
DB 1561 AACATGAGTGAAGCCATFAAAGGGTTCCTGGTCAAGTGTGTGCATACAAATTCCTCTGGC 1620
QY 1621 ACATCTGTGAGACGATCCTTTTAAACTCTCCAGGCCCTTCCCTTTTCAATCCAAGACAAC 1680
DB |||||
DB 1621 ACATCTGTGAGACGATCCTTTTAAACTCTCCAGGCCCTTCCCTTTTCAATCCAAGACAAC 1680
QY 1681 ATCTCATTTCTATGCAACAATTTGTTGTTCTCTCTCTTTTCAATGTCGTTTAAACCTCTGCTA 1740
DB |||||
DB 1681 ATCTCATTTCTATGCAACAATTTGTTGTTCTCTCTCTTTTCAATGTCGTTTAAACCTCTGCTA 1740
QY 1741 ATTTGTCAAGATACAAAAGCAATTTTAGGTATGAAAAGCCAGCTACAGATGTGTACAGGTG 1800
DB |||||
DB 1741 ATTTGTCAAGATACAAAAGCAATTTTAGGTATGAAAAGCCAGCTACAGATGTGTACAGGTG 1800
QY 1801 ACCGCTCTCTCAGATATAGTACTTCTACGTTGATTTTCAGAGATATGAAATATGATCTC 1860
DB |||||
DB 1801 ACCGCTCTCTCAGATATAGTACTTCTACGTTGATTTTCAGAGATATGAAATATGATCTC 1860
QY 1861 AAATGGGAGTTTCCAAGAGAAAATTTAGAGTTTGGAAAGGTACTAGGATCAGGTGCTTTT 1920
DB |||||
DB 1861 AAATGGGAGTTTCCAAGAGAAAATTTAGAGTTTGGAAAGGTACTAGGATCAGGTGCTTTT 1920
QY 1921 GGAAGAGTGTATGAACGCAACAGCTTATGGAATTAGCAAAACAGGAGTCTCAATCCAGGTT 1980
DB |||||
DB 1921 GGAAGAGTGTATGAACGCAACAGCTTATGGAATTAGCAAAACAGGAGTCTCAATCCAGGTT 1980
QY 1981 GCCGTCAAAATGCTGAAAGAAAAGCAGACAGCTCTGAAAGAGAGGCACTCATGTGAGAA 2040
DB |||||
DB 1981 GCCGTCAAAATGCTGAAAGAAAAGCAGACAGCTCTGAAAGAGAGGCACTCATGTGAGAA 2040
QY 2041 CTCAAGATGATGACCCAGCTGGGAAGCACCAGAGAAATTTGTGAACCTCTGGGGCGTGC 2100
DB |||||
DB 2041 CTCAAGATGATGACCCAGCTGGGAAGCACCAGAGAAATTTGTGAACCTCTGGGGCGTGC 2100
QY 2101 ACCTGTGAGSACCAATTTACTTGAATTTTGAATACTGTTGCTATGTTGATCTCTCAAC 2160
DB |||||
DB 2101 ACCTGTGAGSACCAATTTACTTGAATTTTGAATACTGTTGCTATGTTGATCTCTCAAC 2160
QY 2161 TATCTAAGAGTAAAGAGAAAATTTTCAAGGACTTGGACAGAGATTTTCAAGGAAACAC 2220
DB |||||
DB 2161 TATCTAAGAGTAAAGAGAAAATTTTCAAGGACTTGGACAGAGATTTTCAAGGAAACAC 2220
QY 2221 AATTTGAGTTTACCCCACTTTCCAAATCTCAGGGCTTCATGGGAATTCATTTCAC 2280
DB |||||
DB 2221 AATTTGAGTTTACCCCACTTTCCAAATCTCAGGGCTTCATGGGAATTCATTTCAC 2280
QY 2281 GAAGTTCAGATACACCCGGACTCGGATCAATCTCAGGGCTTCATGGGAATTCATTTCAC 2340
DB |||||
DB 2281 GAAGTTCAGATACACCCGGACTCGGATCAATCTCAGGGCTTCATGGGAATTCATTTCAC 2340
QY 2341 TCTGAAGATGAAATTTGAATATGAAAACCAAAAAGGCTGGAAGAGAGGAGGACTTGAAT 2400
DB |||||
DB 2341 TCTGAAGATGAAATTTGAATATGAAAACCAAAAAGGCTGGAAGAGAGGAGGACTTGAAT 2400
QY 2401 GTGCTTACATTTGAGATCTTCTTTGCTTTGCTATCAATCAAGTTGCAAGGAATGGAATTT 2460
DB |||||
DB 2401 GTGCTTACATTTGAGATCTTCTTTGCTTTGCTATCAATCAAGTTGCAAGGAATGGAATTT 2460
QY 2461 CTGGAATTTTAAAGTGTGTGTTTCAAGAGACTGCGCCAGGAACGCTGTTGTCAACCCAC 2520
DB |||||
DB 2461 CTGGAATTTTAAAGTGTGTGTTTCAAGAGACTGCGCCAGGAACGCTGTTGTCAACCCAC 2520
QY 2521 GGGAAAGTGTGAAGATATGTGATTTGGATTTGGCTCGAGATATCATGATGATTTCCAAC 2580

```
Db 2521 GGGAAAGTGTGAAGATATGTGACTTTGGATGGCTCGAGATATCATAGTGATTCACAC 2580
Qy 2581 TATGTTGTTCAGGGGCAATGCCGCTCTCCCTGTAAATGGATGGCCCCGAAAGCCTGTTT 2640
Db 2581 TATGTTGTTCAGGGGCAATGCCGCTCTCCCTGTAAATGGATGGCCCCGAAAGCCTGTTT 2640
Qy 2641 GAAGGCATCTACACCATTAAGAGTGATGTCTGGTCATATGGAATATTAATCTGTGGGAATC 2700
Db 2641 GAAGGCATCTACACCATTAAGAGTGATGTCTGGTCATATGGAATATTAATCTGTGGGAATC 2700
Qy 2701 TTCTCACTTGGTGTGAATCTTACCTGGCATTCGGTGTGATGCTTAATCTCAAACTG 2760
Db 2701 TTCTCACTTGGTGTGAATCTTACCTGGCATTCGGTGTGATGCTTAATCTCAAACTG 2760
Qy 2761 ATTCAAAATGGATTTAAATGGATCAGCCATTTTATGCTACAGAAGAAATATACATTATA 2820
Db 2761 ATTCAAAATGGATTTAAATGGATCAGCCATTTTATGCTACAGAAGAAATATACATTATA 2820
Qy 2821 ATGCAATCTCTGGGCTTTTGAATCAAGGAACGGCCATCTTCCCTAAATTTGACTTCG 2880
Db 2821 ATGCAATCTCTGGGCTTTTGAATCAAGGAACGGCCATCTTCCCTAAATTTGACTTCG 2880
Qy 2881 TTTTATGATCTCAGCTGGCAGATGCGAAGAGCGATGATCAGAATGTGGATGGCCGT 2940
Db 2881 TTTTATGATCTCAGCTGGCAGATGCGAAGAGCGATGATCAGAATGTGGATGGCCGT 2940
Qy 2941 GTTTCGGAATGTCCTCACACCTTACCAAAACAGGCGACCTTTTCAGCAGAGATGGATTG 3000
Db 2941 GTTTCGGAATGTCCTCACACCTTACCAAAACAGGCGACCTTTTCAGCAGAGATGGATTG 3000
Qy 3001 GGGCTACTCTCTCCGAGGCTCAGGTGGAAGATCTGTAGAGGAACAATTTAGTTTTAAG 3060
Db 3001 GGGCTACTCTCTCCGAGGCTCAGGTGGAAGATCTGTAGAGGAACAATTTAGTTTTAAG 3060
Qy 3061 ACTTCATCCCTCCACCTATCCCTAACAGGCTGTAGATTACCAAAACAGATTAATTCAT 3120
Db 3061 ACTTCATCCCTCCACCTATCCCTAACAGGCTGTAGATTACCAAAACAGATTAATTCAT 3120
Qy 3121 CACTAAAAGAAAATCTATTATCAACTGCTCTCCACAGACCTTTTCTAGAACCGCTCT 3180
Db 3121 CACTAAAAGAAAATCTATTATCAACTGCTCTCCACAGACCTTTTCTAGAACCGCTCT 3180
Qy 3181 GCGTTTACTCTGTTTCAAGAGGACTTTTCTAAAATCAAAATCATCTCTGTACAAAGCAG 3240
Db 3181 GCGTTTACTCTGTTTCAAGAGGACTTTTCTAAAATCAAAATCATCTCTGTACAAAGCAG 3240
Qy 3241 GAGGAGCTGATATGAATTTTATGGAGCATTTGATCTGCATCCAGGCTTCTCAGGCG 3300
Db 3241 GAGGAGCTGATATGAATTTTATGGAGCATTTGATCTGCATCCAGGCTTCTCAGGCG 3300
Qy 3301 GCTTGAAGTGAATCTGTACCTGAAGTACAGTATATCTTGTAAATACATAAAACAAAGC 3360
Db 3301 GCTTGAAGTGAATCTGTACCTGAAGTACAGTATATCTTGTAAATACATAAAACAAAGC 3360
Qy 3361 ATTTTGTCTAAGGAGAGCTAATATGATTTTAAAGTCTATGTTTTTAAATTAATATGAAA 3420
Db 3361 ATTTTGTCTAAGGAGAGCTAATATGATTTTAAAGTCTATGTTTTTAAATTAATATGAAA 3420
Qy 3421 TTTTTCAGCTATTAGTATATATTTTATGGTGGGAATAAAATTTCTACTACAGAAAAA 3480
Db 3421 TTTTTCAGCTATTAGTATATATTTTATGGTGGGAATAAAATTTCTACTACAGAAAAA 3480
Qy 3481 AAAAAAAAAAAAAAAAAAAAAA 3501
Db 3481 AAAAAAAAAAAAAAAAAAAAAA 3501
```

RESULT 7

PCT-US92-09893-3

; Sequence 3, Application PC/TUS9209893

; GENERAL INFORMATION:

; APPLICANT: Lemischka, Ihor R.

```
; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL
; TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ImClone Systems Incorporated
; STREET: 180 Varick Street
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10014
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/09893
; FILING DATE: 19921116
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Feit, Irving N.
; REGISTRATION NUMBER: 28,601
; REFERENCE/DOCKET NUMBER: LEM-3-7PT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-645-1405
; TELEFAX: 212-645-2054
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3501 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: double
; TOPOLOGY: linear
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 58..3039
; FEATURE:
; NAME/KEY: mat peptide
; LOCATION: 139..3036
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 58..138
; PCT-US92-09893-3
```

```
Query Match 100.0%; Score 3501; DB 5; Length 3501;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGAGGGCGCATCCGAGGGCTGGGGCGGCGCCCTGGGGAGACCCCGGGCTCCGGAGGCCATG 60
Db 1 CGAGGGCGCATCCGAGGGCTGGGGCGGCGCCCTGGGGAGACCCCGGGCTCCGGAGGCCATG 60

Qy 61 CCGCGTTGGCGCGCGACGCGGGCACCGTGCCTGCTGCTGTTTTTCTGCAATGATA 120
Db 61 CCGCGTTGGCGCGCGACGCGGGCACCGTGCCTGCTGCTGTTTTTCTGCAATGATA 120

Qy 121 TTTGGGACTATTACAAATCAAGATCTGCCTGTGATCAAGTGTGTTTTTAATCATATAAG 180
Db 121 TTTGGGACTATTACAAATCAAGATCTGCCTGTGATCAAGTGTGTTTTTAATCATATAAG 180

Qy 181 AACAAATGATTCATCAGTGGGAGTCAATCATATCCCATGTTATCAGAAATCCCGGAA 240
Db 181 AACAAATGATTCATCAGTGGGAGTCAATCATATCCCATGTTATCAGAAATCCCGGAA 240

Qy 241 GACCTCGGGTGTGCGTTTGAGACCCAGAGCTCAGGGAAGTGTACGAAGCTGCCGCTGTG 300
Db 241 GACCTCGGGTGTGCGTTTGAGACCCAGAGCTCAGGGAAGTGTACGAAGCTGCCGCTGTG 300

Qy 301 GAAGTGGATGATCTGTTTCCATCACACTGCAAGTGTGTCGATGCCCGGAGGACATT 360
Db 301 GAAGTGGATGATCTGTTTCCATCACACTGCAAGTGTGTCGATGCCCGGAGGACATT 360
```

Db 301 GAAGTGGATGTATCTGCTTCCATCACACTGCAAGTGTGGTCGATGCCCGGGAACATT 360
Qy TCCTGTCTCTGGGTCTTTAAGCAGACAGCTCCCTGAAATGCCAGGCCACATTTTGAATTACAA 420
Db TCCTGTCTCTGGGTCTTTAAGCAGACAGCTCCCTGAAATGCCAGGCCACATTTTGAATTACAA 420
Qy 421 AACAGAGAGTGTCTCATGGTCAATTTTGAATAATGACAGAAACCCAGCTGGAGATAC 480
Db AACAGAGAGTGTCTCATGGTCAATTTTGAATAATGACAGAAACCCAGCTGGAGATAC 480
Qy 481 CTACTTTTATTCAGAGTGAAGCTACCAATTTACCAATATTTGTTTACAGTGAATTAAGA 540
Db CTACTTTTATTCAGAGTGAAGCTACCAATTTACCAATATTTGTTTACAGTGAATTAAGA 540
Qy 541 AATACCCTGCTTTACACATTTAAGAGACCTTACCTTTTGAATAATGGAATAACCCAGAGGCC 600
Db AATACCCTGCTTTACACATTTAAGAGACCTTACCTTTTGAATAATGGAATAACCCAGAGGCC 600
Qy 601 CTGCTGTGCATATCTGAGAGCTTCCAGAGCCGATCGTGGAAATGGGTGCTTTTGGGATTC 660
Db CTGCTGTGCATATCTGAGAGCTTCCAGAGCCGATCGTGGAAATGGGTGCTTTTGGGATTC 660
Qy 661 CAGGGGAAAGCTGTAAGAGAAAGTCCAGCTGTTTAAAGAGGGAATAAGTGCCT 720
Db CAGGGGAAAGCTGTAAGAGAAAGTCCAGCTGTTTAAAGAGGGAATAAGTGCCT 720
Qy 721 CATGAATTTATTCGGAGCGACATAAGCTGCTGTCAGCAAGATGAATGCGGAGGGAATGC 780
Db CATGAATTTATTCGGAGCGACATAAGCTGCTGTCAGCAAGATGAATGCGGAGGGAATGC 780
Qy 781 ACCAGGCTGTTCAATATAGATCTAAATCAAACTCCTCAGACCAACATTTGCCACAAATTTT 840
Db ACCAGGCTGTTCAATATAGATCTAAATCAAACTCCTCAGACCAACATTTGCCACAAATTTT 840
Qy 841 CTTAAAGTAGGGAAACCTTTATGGATAGGTGCAAGCTGTTTCATGTGAACCAATGGAATTC 900
Db CTTAAAGTAGGGAAACCTTTATGGATAGGTGCAAGCTGTTTCATGTGAACCAATGGAATTC 900
Qy 901 GGGCTCACCTGGGAAATTTAGAAAACAAAGCACTCGAGGAGGCACTACTTTGAGATGAGT 960
Db GGGCTCACCTGGGAAATTTAGAAAACAAAGCACTCGAGGAGGCACTACTTTGAGATGAGT 960
Qy 961 ACCTATTCAACAAACAGAACTATGATACGGATCTGTTTCTGCTTTGATCATCATGCGCA 1020
Db ACCTATTCAACAAACAGAACTATGATACGGATCTGTTTCTGCTTTGATCATCATGCGCA 1020
Qy 1021 AGAAACGACACCGGATACACTGTTTCTCTTCAAAGCATCCAGTCAATCAGCTTTG 1080
Db AGAAACGACACCGGATACACTGTTTCTCTTCAAAGCATCCAGTCAATCAGCTTTG 1080
Qy 1081 GTTACCATCGTAGAAAGGATTTATTAATGCTACCAATTTCAAGTGAAGATTTAGAAAT 1140
Db GTTACCATCGTAGAAAGGATTTATTAATGCTACCAATTTCAAGTGAAGATTTAGAAAT 1140
Qy 1141 GACCAATATGAAGATTTTCTGTTTCTGTCAGGTTTAAAGCCCTACCCACAAATCAGATGT 1200
Db GACCAATATGAAGATTTTCTGTTTCTGTCAGGTTTAAAGCCCTACCCACAAATCAGATGT 1200
Qy 1201 ACGTGGACCTTCTCTGAAAATCATTTTCTGTCAGCAAAAGGCTTTGATAACGGATAC 1260
Db ACGTGGACCTTCTCTGAAAATCATTTTCTGTCAGGTTTAAAGCCCTACCCACAAATCAGATGT 1260
Qy 1261 AGCATATCCAGTTTTCGAAATCATAGCACCGCAGGAGAAATATATTTCCATGAGAA 1320
Db AGCATATCCAGTTTTCGAAATCATAGCACCGCAGGAGAAATATATTTCCATGAGAA 1320
Qy 1321 AATGATGATGCCAAATTTACCAAAATGTTTACCGCTGAATATAGAGGAACCTCAAGT 1380
Db AATGATGATGCCAAATTTACCAAAATGTTTACCGCTGAATATAGAGGAACCTCAAGT 1380
Qy 1381 CTCGCAAGAGCATCGGCAAGTCAAGGCTCTGTTTCTCGGATGGATACCCATACCATCT 1440
Db CTCGCAAGAGCATCGGCAAGTCAAGGCTCTGTTTCTCGGATGGATACCCATACCATCT 1440

Qy 1441 TGGACCTGGGAAGAGTGTTCAGACAAGTCTCCCACTCCACAGAGAGATCACAGAAGA 1500
Db TGGACCTGGGAAGAGTGTTCAGACAAGTCTCCCACTCCACAGAGAGATCACAGAAGA 1500
Qy 1501 GTCTGGAATAGAAAGGCTTAACAGAAAAAGTGTTCGACAGTGGGTGTCGAGCAGTACTCTA 1560
Db GTCTGGAATAGAAAGGCTTAACAGAAAAAGTGTTCGACAGTGGGTGTCGAGCAGTACTCTA 1560
Qy 1561 AACATGATGAAGCCATTAAGAGGTTCTCTGGTCAAGTGTCTGTCATACAAATCCCTTGGC 1620
Db AACATGATGAAGCCATTAAGAGGTTCTCTGGTCAAGTGTCTGTCATACAAATCCCTTGGC 1620
Qy 1621 ACATCTTGTGAGAGATCCCTTTTAAACTCTCCAGGCCCTTCCCTTTTCAATCAAGACAA 1680
Db ACATCTTGTGAGAGATCCCTTTTAAACTCTCCAGGCCCTTCCCTTTTCAATCAAGACAA 1680
Qy 1681 ATCTCATTTATGCAACAAATTTGGTGTGTTGCTCTCTCTTCAATGCTGTTTAAACCTGCTA 1740
Db ATCTCATTTATGCAACAAATTTGGTGTGTTGCTCTCTCTTCAATGCTGTTTAAACCTGCTA 1740
Qy 1741 ATTTGTCAAGTACAAAAAGCAATTTAGGTATGAAAGCCAGCTACAGATGTTACAGGTG 1800
Db ATTTGTCAAGTACAAAAAGCAATTTAGGTATGAAAGCCAGCTACAGATGTTACAGGTG 1800
Qy 1801 ACCGGCTCCTCAGATATGATGATCTCTACGTTGATTTTCAAGAAATATGAATATGATCTC 1860
Db ACCGGCTCCTCAGATATGATGATCTCTCTACGTTGATTTTCAAGAAATATGAATATGATCTC 1860
Qy 1861 AATGGGAGTTTCCAGAGAAAAATTTAGAGTTTGGGAAGTACTAGGATCAGGTGCTTTT 1920
Db AATGGGAGTTTCCAGAGAAAAATTTAGAGTTTGGGAAGTACTAGGATCAGGTGCTTTT 1920
Qy 1921 GGAAGAGTGAAGAACGCAACAGCTTATGAAATTAGCAAAAACAGAGTCTCAATCCAGGTT 1980
Db GGAAGAGTGAAGAACGCAACAGCTTATGAAATTAGCAAAAACAGAGTCTCAATCCAGGTT 1980
Qy 1981 GCGTCAAAATGCTGAAAGAAAAAGCAGACAGCTCTGAAAGAGAGGCACTCATGTCAGAA 2040
Db GCGTCAAAATGCTGAAAGAAAAAGCAGACAGCTCTGAAAGAGAGGCACTCATGTCAGAA 2040
Qy 2041 CTCAGATGATGACCCAGCTGGGAAGCCAGAGATTTTGAACCTGCTGGGGGCGTGC 2100
Db CTCAGATGATGACCCAGCTGGGAAGCCAGAGATTTTGAACCTGCTGGGGGCGTGC 2100
Qy 2101 ACACTGTGAGGACCAATTTTACTGTTTGAATACGTTTGTGCTATGTTGATCTTCTCAAC 2160
Db ACACTGTGAGGACCAATTTTACTGTTTGAATACGTTTGTGCTATGTTGATCTTCTCAAC 2160
Qy 2161 TATCTAAGAGTAAAAAGAGAAAAATTTTCAAGGACTTTGGAAGAGATTTTCAAGGAACAC 2220
Db TATCTAAGAGTAAAAAGAGAAAAATTTTCAAGGACTTTGGAAGAGATTTTCAAGGAACAC 2220
Qy 2221 AATTTTCAGTTTTCACCCACTTTCCATCAGATCAGATCAGATCAGATCAGATCAGATCAG 2280
Db AATTTTCAGTTTTCACCCACTTTCCATCAGATCAGATCAGATCAGATCAGATCAGATCAG 2280
Qy 2281 GAAGTTTCAAGATACACCGGACTCGGATCAAACTCTCAGGGCTTCAATGGGATTTCAATTC 2340
Db GAAGTTTCAAGATACACCGGACTCGGATCAAACTCTCAGGGCTTCAATGGGATTTCAATTC 2340
Qy 2341 TCTGAAGATGAATTTGAATATGAAAAACCAAAAGGCTTGAAGAGAGGAGGACTTTGAAT 2400
Db TCTGAAGATGAATTTGAATATGAAAAACCAAAAGGCTTGAAGAGAGGAGGACTTTGAAT 2400
Qy 2401 GTGCTTACATTTGAAGATCTTCTTGTGCTTTCAGATCAAGTTTCCAAAGGAAATGGAAT 2460
Db GTGCTTACATTTGAAGATCTTCTTGTGCTTTCAGATCAAGTTTCCAAAGGAAATGGAAT 2460
Qy 2461 CTGGAATTTAAGTCGTGTTTCAAGAGACCTGGCCGCGGAGGAACTGCTTGTGTCACCCAC 2520
Db CTGGAATTTAAGTCGTGTTTCAAGAGACCTGGCCGCGGAGGAACTGCTTGTGTCACCCAC 2520

421 AACAGAGGAGTTGTTTCCATGGTCAATTTGAAATGACAGAAACCCCAAGCTGAGAAATAC 480
Db
421 AACAGAGGAGTTGTTTCCATGGTCAATTTGAAATGACAGAAACCCCAAGCTGAGAAATAC 480
Qy
481 CTACTTTTATTCAGAGTGAAGCTACCAATTACCAATATTTGTTTACAGTGAATTAAGA 540
Db
481 CTACTTTTATTCAGAGTGAAGCTACCAATTACCAATATTTGTTTACAGTGAATTAAGA 540
Qy
541 AATACCTCTGTTTACACATTAGAAAGACCTTACTTTAGAAATGAAACACAGGACGC 600
Db
541 AATACCTCTGTTTACACATTAGAAAGACCTTACTTTAGAAATGAAACACAGGACGC 600
Qy
601 CTGGTCTGCATATCTGAGAGCGTTCCAGAGCCGATCGTGAATGGGTGCTTTCCGATTCA 660
Db
601 CTGGTCTGCATATCTGAGAGCGTTCCAGAGCCGATCGTGAATGGGTGCTTTCCGATTCA 660
Qy
661 CAGGGGAAAGCTGTTAAAGAAAGTCCAGCTGTTGTTTAAAGAGGAGGAAAGTGCTT 720
Db
661 CAGGGGAAAGCTGTTAAAGAAAGTCCAGCTGTTGTTTAAAGAGGAGGAAAGTGCTT 720
Qy
721 CATGAATTTTGGAGCGACATAGGTGCTGTCAGAAATGAATGAATGGGCGAGGATGC 780
Db
721 CATGAATTTTGGAGCGACATAGGTGCTGTCAGAAATGAATGAATGGGCGAGGATGC 780
Qy
781 ACCAGGCTGTTCAACAATAGATCTAAATCAAACTCCTCAGACACACATTTGCCAATTTATTT 840
Db
781 ACCAGGCTGTTCAACAATAGATCTAAATCAAACTCCTCAGACACACATTTGCCAATTTATTT 840
Qy
841 CTTAAAGTAGGGGAAACCTTTATGGATAAGGTGCAAAAGCTGTTTCATGTGAACCATGGATTC 900
Db
841 CTTAAAGTAGGGGAAACCTTTATGGATAAGGTGCAAAAGCTGTTTCATGTGAACCATGGATTC 900
Qy
901 GGGCTCACCTGGGAATTAGAAACAAAGCACTCGAGGAGGCAACTPCTTGAGATGAGT 960
Db
901 GGGCTCACCTGGGAATTAGAAACAAAGCACTCGAGGAGGCAACTPCTTGAGATGAGT 960
Qy
961 ACCTATTCAACAAACAGAACTATGATACGGATCTGTTGCTTTGTTATCATCAGTGGGA 1020
Db
961 ACCTATTCAACAAACAGAACTATGATACGGATCTGTTGCTTTGTTATCATCAGTGGGA 1020
Qy
1021 AGAAGCAGACCCGGATCTACACTTGTCTCTTCAAGCATCCCAAGTCAATCAGCTTNG 1080
Db
1021 AGAAGCAGACCCGGATCTACACTTGTCTCTTCAAGCATCCCAAGTCAATCAGCTTNG 1080
Qy
1081 GTTACCATCGTAGGAAGGGAATTAATAATGCTACCAATTCAGTGAAGATTAATAATTT 1140
Db
1081 GTTACCATCGTAGGAAGGGAATTAATAATGCTACCAATTCAGTGAAGATTAATAATTT 1140
Qy
1141 GACCAATATGAAGATTTGTTTCTGTCAGGTTTAAAGCCTACCCCAAAATCAGATGT 1200
Db
1141 GACCAATATGAAGATTTGTTTCTGTCAGGTTTAAAGCCTACCCCAAAATCAGATGT 1200
Qy
1201 AGCTGGACCTCTCTCGAAAATCATTTCTCTGTGAGCAAAAGGCTTGTGATAACGGATAC 1260
Db
1201 AGCTGGACCTCTCTCGAAAATCATTTCTCTGTGAGCAAAAGGCTTGTGATAACGGATAC 1260
Qy
1261 AGCATATCAAGTTTGGCAATCATAGCACACAGCCAGGAGATATATATTCATGCGAGAA 1320
Db
1261 AGCATATCAAGTTTGGCAATCATAGCACACAGCCAGGAGATATATATTCATGCGAGAA 1320
Qy
1321 AATGATGATGCCAATTTACCAAAATGTTTACGCTGAATATTAAGAGGAAACCTCAAGTG 1380
Db
1321 AATGATGATGCCAATTTACCAAAATGTTTACGCTGAATATTAAGAGGAAACCTCAAGTG 1380
Qy
1381 CTCGAGAGCATCGGCAAGTCTCCCAAGTCTCCCAACTGCAAGAGAGATCAAGAGGA 1440
Db
1381 CTCGAGAGCATCGGCAAGTCTCCCAAGTCTCCCAACTGCAAGAGAGATCAAGAGGA 1440
Qy
1441 TGGACCTGGAAGAGTGTTCAGACAAAGTCTCCCAAGTCTCCCAACTGCAAGAGAGATCAAGAGGA 1500
Db
1441 TGGACCTGGAAGAGTGTTCAGACAAAGTCTCCCAAGTCTCCCAACTGCAAGAGAGATCAAGAGGA 1500

1501 GTCTGGAATAGAAAGCGCTAAACAGAAAGTGTGTCAGTGGGTGTCGAGCAGTACTCTA 1560
Db
1501 GTCTGGAATAGAAAGCGCTAAACAGAAAGTGTGTCAGTGGGTGTCGAGCAGTACTCTA 1560
Qy
1561 AACATGAGTGAAGCCATAAAGGGTTCCTGGTCAAGTGTGTCATACAATTTCCCTTGGC 1620
Db
1561 AACATGAGTGAAGCCATAAAGGGTTCCTGGTCAAGTGTGTCATACAATTTCCCTTGGC 1620
Qy
1621 ACATCTTGTGAGACGATCCCTTTAACTCTCAGGCCCCCTTCCCTTTTATCCAGACAAC 1680
Db
1621 ACATCTTGTGAGACGATCCCTTTAACTCTCAGGCCCCCTTCCCTTTTATCCAGACAAC 1680
Qy
1681 ATCTCATTTCTATGCAACAATTTGTTGTTCTCTCTCTCATTTGTCGTTTTTAAACCTGCTA 1740
Db
1681 ATCTCATTTCTATGCAACAATTTGTTGTTCTCTCTCTCATTTGTCGTTTTTAAACCTGCTA 1740
Qy
1741 ATTTGTCAAGTACAAAAGCAATTTAGTATGAAAGCAGCTACAGATGGTACAGGTG 1800
Db
1741 ATTTGTCAAGTACAAAAGCAATTTAGTATGAAAGCAGCTACAGATGGTACAGGTG 1800
Qy
1801 ACCGGCTCTCAGATAATGAGTACTTCTACGTTGATTTTACAGAGATATGAATATGATCTC 1860
Db
1801 ACCGGCTCTCAGATAATGAGTACTTCTACGTTGATTTTACAGAGATATGAATATGATCTC 1860
Qy
1861 AAATGGGAGTTTCCAGAGAAATTTAGAGTTTGGGAGGTACTAGGATCAGGTGCTTTT 1920
Db
1861 AAATGGGAGTTTCCAGAGAAATTTAGAGTTTGGGAGGTACTAGGATCAGGTGCTTTT 1920
Qy
1921 GGAAAGTGTATGAACGCAACAGCTTATGAAATTTAGCAAAACAGGAGTCTCAATCCAGTT 1980
Db
1921 GGAAAGTGTATGAACGCAACAGCTTATGAAATTTAGCAAAACAGGAGTCTCAATCCAGTT 1980
Qy
1981 GCCGTCAAAATGCTCAAGAAAGCAAGCAGCTCTGAAAGAGAGGCACTCATGTGAGAA 2040
Db
1981 GCCGTCAAAATGCTCAAGAAAGCAAGCAGCTCTGAAAGAGAGGCACTCATGTGAGAA 2040
Qy
2041 CTCAGATGATGACCCAGCTGGGAGGCAAGCAATATTTGTGAACCTGCTGGGGGGGTGC 2100
Db
2041 CTCAGATGATGACCCAGCTGGGAGGCAAGCAATATTTGTGAACCTGCTGGGGGGGTGC 2100
Qy
2101 ACACGTGTGAGGACCAATTTTACTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTT 2160
Db
2101 ACACGTGTGAGGACCAATTTTACTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTT 2160
Qy
2161 TATCTAAGAGTAAAGAGAAATTTTCAAGGATTTGACAGAGATTTTCAAGGAAACAC 2220
Db
2161 TATCTAAGAGTAAAGAGAAATTTTCAAGGATTTGACAGAGATTTTCAAGGAAACAC 2220
Qy
2221 AATTTCAAGTTTACCCCACTTTTCAATCAATCCAAATTTCCAGATGCTGTTCAAGA 2280
Db
2221 AATTTCAAGTTTACCCCACTTTTCAATCAATCCAAATTTCCAGATGCTGTTCAAGA 2280
Qy
2281 GAAGTTCAAGTACACCCGAGCTCGGATCAATTTTCAAGGCTTTCATGGGAAATTTCAATTCAC 2340
Db
2281 GAAGTTCAAGTACACCCGAGCTCGGATCAATTTTCAAGGCTTTCATGGGAAATTTCAATTCAC 2340
Qy
2341 TCTGAAGATGAATTAAGATTAAGAAACCAAAAGGCTTGGAGAGAGAGGAGGACTTGAAT 2400
Db
2341 TCTGAAGATGAATTAAGATTAAGAAACCAAAAGGCTTGGAGAGAGAGGAGGACTTGAAT 2400
Qy
2401 GTGCTTCAATTTGAAGATCTTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTT 2460
Db
2401 GTGCTTCAATTTGAAGATCTTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTT 2460
Qy
2461 CTGGAATTTAAGTGTGTTTCAAGACCTTGGCCGCGAGGAACTGCTGTTGTCACCCAC 2520
Db
2461 CTGGAATTTAAGTGTGTTTCAAGACCTTGGCCGCGAGGAACTGCTGTTGTCACCCAC 2520
Qy
2521 GGGAAAGTGTGAAGATATGATGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTT 2580
Db
2521 GGGAAAGTGTGAAGATATGATGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTT 2580
Qy
2581 TATGTTGTGAGGGCAATGCGCTGCTGCTGTTAAATGGATGGCGCCCGGAAAGGCTGTTT 2640

Db 2581 TATGTTGTCAGGGCAATGCCGCTGCTGCTGTAATGATGGCCCCGAAAGCCCTGTTT 2640
Qy 2641 GAAGGCATCTACACCATTAAGAGTGATGCTGGTCATATGGAATATTACTGTGGAAATC 2700
Db 2641 GAAGGCATCTACACCATTAAGAGTGATGCTGGTCATATGGAATATTACTGTGGAAATC 2700
Qy 2701 TTCTCATTGTTGTAATCCTTACCTGGCATTCGGTTGATGCTAACTTCTACAAATC 2760
Db 2701 TTCTCATTGTTGTAATCCTTACCTGGCATTCGGTTGATGCTAACTTCTACAAATC 2760
Qy 2761 ATTCAAATGATTTAAATGATCAGCATTTATGCTACAGAAAGAAATATACATTATA 2820
Db 2761 ATTCAAATGATTTAAATGATCAGCATTTATGCTACAGAAAGAAATATACATTATA 2820
Qy 2821 ATGCAATCCTGCTGGGCTTTTGACTCAAGGAAACGGCCATCCTTCCCTAAATTGACTTCG 2880
Db 2821 ATGCAATCCTGCTGGGCTTTTGACTCAAGGAAACGGCCATCCTTCCCTAAATTGACTTCG 2880
Qy 2881 TTTTAGATGTCAGCTGGCAGATCAGAGGAGGATGATCAGAAATGATGATGATGATGATG 2940
Db 2881 TTTTAGATGTCAGCTGGCAGATCAGAGGAGGATGATCAGAAATGATGATGATGATGATG 2940
Qy 2941 GTTTCGGAATGCTCCTCACACCTTACCAAAACAGCGACCTTTTCAGCAGAGAGATGATTTG 3000
Db 2941 GTTTCGGAATGCTCCTCACACCTTACCAAAACAGCGACCTTTTCAGCAGAGAGATGATTTG 3000
Qy 3001 GGGCTACTCTCTCCGAGGCTCAGGTCGAAGATTCGTAGAGAAACAATTTAGTTTAAAG 3060
Db 3001 GGGCTACTCTCTCCGAGGCTCAGGTCGAAGATTCGTAGAGAAACAATTTAGTTTAAAG 3060
Qy 3061 ACTTCATCCTCCACCTATCCCTAACAGGCTGATGATTAACAAAACAAGATTAATTTAT 3120
Db 3061 ACTTCATCCTCCACCTATCCCTAACAGGCTGATGATTAACAAAACAAGATTAATTTAT 3120
Qy 3121 CACTAAAAGAAATCTATTATCAATGCTGCTTCCAGCAGCTTTTCTAGAGCGGCT 3180
Db 3121 CACTAAAAGAAATCTATTATCAATGCTGCTTCCAGCAGCTTTTCTAGAGCGGCT 3180
Qy 3181 GGGTTTACTCTGTTTCAAGGGGACTTTTGTAAATCAATCATCTCTGTCACAGGCG 3240
Db 3181 GGGTTTACTCTGTTTCAAGGGGACTTTTGTAAATCAATCATCTCTGTCACAGGCG 3240
Qy 3241 GAGGAGTGATTAATGAATTTATGAGCATTTGATGATCAATCAATCAATCAATCAATCA 3300
Db 3241 GAGGAGTGATTAATGAATTTATGAGCATTTGATGATCAATCAATCAATCAATCAATCA 3300
Qy 3301 GCTTCAGTGATTTGTTACCTGACGTACAGTATTTCTGTAATACATAAAACAAGC 3360
Db 3301 GCTTCAGTGATTTGTTACCTGACGTACAGTATTTCTGTAATACATAAAACAAGC 3360
Qy 3361 ATTTTGTCTAAGGAGAGCTTAATATGATTTTAAAGTCTATGTTTAAATAATATATATA 3420
Db 3361 ATTTTGTCTAAGGAGAGCTTAATATGATTTTAAAGTCTATGTTTAAATAATATATATA 3420
Qy 3421 TTTTTCAGCTATTAGTGATATATTTTATGGTGGGAAATAAATTTCTACTACAG 3475
Db 3421 TTTTTCAGCTATTAGTGATATATTTTATGGTGGGAAATAAATTTCTACTACAG 3475

RESULT 9

US-08-434-878-3
; Sequence 3, Application US/08434878
; Patent No. 5997865

GENERAL INFORMATION:

; APPLICANT: Bennett, Brian D.
; APPLICANT: Broz, Susan D.
; APPLICANT: Matthews, William
; APPLICANT: Zeigler, Francis C.
; TITLE OF INVENTION: AGONIST ANTIBODIES AGAINST THE FLK2/FLT3 RECEPTOR AND USES THEREIN
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.

STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: patin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/434,878
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Hasak, Janet E.
REGISTRATION NUMBER: 28,616
REFERENCE/DOCKET NUMBER: 879
TELEPHONE: 415/225-1896
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 3475 bases
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-434-878-3

Query Match 99.3%; Score 3475; DB 2; Length 3475;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGAGCGGCATCCGAGGGCTGGCGCGGCGCCCTGGGGGACCCCGGGCTCCGAGGCGCATG 60
Db 1 CGAGCGGCATCCGAGGGCTGGCGCGGCGCCCTGGGGGACCCCGGGCTCCGAGGCGCATG 60
Qy 61 CGCGGTTGGCGCGCGCGGCGGCGCCCTGGCGGCTGCTGTTGTTTCTGCAATGATA 120
Db 61 CGCGGTTGGCGCGCGCGGCGGCGCCCTGGCGGCTGCTGTTGTTTCTGCAATGATA 120
Qy 121 TTTGGGACTATTACAAATCAAGATCTGCTGTGATCAAGTGTGTTTAAATCAATCAATAAG 180
Db 121 TTTGGGACTATTACAAATCAAGATCTGCTGTGATCAAGTGTGTTTAAATCAATCAATAAG 180
Qy 181 AACATGATTCATCAGTGGGGAAGTCATCATATCCATGGTATCAGAAATCCCGGAA 240
Db 181 AACATGATTCATCAGTGGGGAAGTCATCATATCCATGGTATCAGAAATCCCGGAA 240
Qy 241 GACCTCGGCTGTGCTTCAGACCCCGAGAGCTCAGGAGCAGTGTACGAGCTCCCGCTGTG 300
Db 241 GACCTCGGCTGTGCTTCAGACCCCGAGAGCTCAGGAGCAGTGTACGAGCTCCCGCTGTG 300
Qy 301 GAAAGTGATGATCTGCTTCCATCACTGCAAGTGTGCTGCTGATGCCCCAGGGAACATT 360
Db 301 GAAAGTGATGATCTGCTTCCATCACTGCAAGTGTGCTGCTGATGCCCCAGGGAACATT 360
Qy 361 TCCTCTCTCTGGGCTTTTAAAGCAGACAGCTCCCTGAAATTCGACCCCAATTTGATTACAA 420
Db 361 TCCTCTCTCTGGGCTTTTAAAGCAGACAGCTCCCTGAAATTCGACCCCAATTTGATTACAA 420
Qy 421 AACAGAGGAGTTGTTTCCATGCTGCTTGAATAATGACAGAAACCCAGCTCGGAGATAC 480
Db 421 AACAGAGGAGTTGTTTCCATGCTGCTTGAATAATGACAGAAACCCAGCTCGGAGATAC 480
Qy 481 CTACTTTTATTACAGGTGAAGCTACCAATATACAAATATTGTTTACAGTGAAGTAAAGA 540
Db 481 CTACTTTTATTACAGGTGAAGCTACCAATATACAAATATTGTTTACAGTGAAGTAAAGA 540

QY 541 AATACCCTGCTTTACATTAAGAAGACCTTTACTTTAGAAAAATGGAAGAACCCAGACGCC 600
DB 541 AATACCCTGCTTTACATTAAGAAGACCTTTACTTTAGAAAAATGGAAGAACCCAGACGCC 600
QY 601 CTGCTCTGCATPATCTGAGAGCGTTCCAGAGCGGATCGTGGAAATGGTGCTTTTGGCATCA 660
DB 601 CTGCTCTGCATPATCTGAGAGCGTTCCAGAGCGGATCGTGGAAATGGTGCTTTTGGCATCA 660
QY 661 CAGGGGAAGCTGTAAAGAAAGATCCAGCTGTTTAAAGAGGAGGAAAGTGCCTT 720
DB 661 CAGGGGAAGCTGTAAAGAAAGATCCAGCTGTTTAAAGAGGAGGAAAGTGCCTT 720
QY 721 CATGAATTTATTTGGGACGACATTAAGGTGCTGCGCAAAATGAATCGGCAGGGAATGC 780
DB 721 CATGAATTTATTTGGGACGACATTAAGGTGCTGCGCAAAATGAATCGGCAGGGAATGC 780
QY 781 ACCAGGCTGTTTCAATATAGATCTTAATCAAACTCCTCAGACCACATTCGCCAATTAFTT 840
DB 781 ACCAGGCTGTTTCAATATAGATCTTAATCAAACTCCTCAGACCACATTCGCCAATTAFTT 840
QY 841 CTTAAAGTAGGGGAACCTTTATGATAGGTGCAAGCTGTTTCATGTGAACCATGGATTC 900
DB 841 CTTAAAGTAGGGGAACCTTTATGATAGGTGCAAGCTGTTTCATGTGAACCATGGATTC 900
QY 901 GGGCTCACCTCGGGAATTAGAAAACAAAGCACTCGAGGAGGCGCACTACTTTGAGATGAT 960
DB 901 GGGCTCACCTCGGGAATTAGAAAACAAAGCACTCGAGGAGGCGCACTACTTTGAGATGAT 960
QY 961 ACCTATTCAACAAACAGACATATGATACGGATTCCTGTTTGTATCATCAGTGSCA 1020
DB 961 ACCTATTCAACAAACAGACATATGATACGGATTCCTGTTTGTATCATCAGTGSCA 1020
QY 1021 AGAAACGACACCGGATACACATTTGTTCTCTTCAAGAGATCCAGTCAATCAGCTTTG 1080
DB 1021 AGAAACGACACCGGATACACATTTGTTCTCTTCAAGAGATCCAGTCAATCAGCTTTG 1080
QY 1081 GTTACCATCTGAGAAAGGATTTATAAATGCTACCAATTCAGTGAAGATTTGAAAT 1140
DB 1081 GTTACCATCTGAGAAAGGATTTATAAATGCTACCAATTCAGTGAAGATTTGAAAT 1140
QY 1141 GACCAATATGAAGAGTTTGTGTTTCTGTCAGGTTTAAAGCTACCCACAAATCAGATGT 1200
DB 1141 GACCAATATGAAGAGTTTGTGTTTCTGTCAGGTTTAAAGCTACCCACAAATCAGATGT 1200
QY 1201 ACGTGGACCTTCTCGGAAATCAATTTCTGTGAGCAAAAGGCTTTGATAACGGATAC 1260
DB 1201 ACGTGGACCTTCTCGGAAATCAATTTCTGTGAGCAAAAGGCTTTGATAACGGATAC 1260
QY 1261 AGCATATCCAAGTTTTCGAATCATAAGCACAGCCAGAGGAATATATATTCATGCGAG 1320
DB 1261 AGCATATCCAAGTTTTCGAATCATAAGCACAGCCAGAGGAATATATATTCATGCGAG 1320
QY 1321 AATCATGATGCCAATTTTACCAAAATGTTTACGCTGTAATATAAGAGGAACCTCAAGTG 1380
DB 1321 AATCATGATGCCAATTTTACCAAAATGTTTACGCTGTAATATAAGAGGAACCTCAAGTG 1380
QY 1381 CTCGAGAAGCATCGGCAAGTCAGGCGTCTGTTTCTCGGATGGATACCCATTAACATCT 1440
DB 1381 CTCGAGAAGCATCGGCAAGTCAGGCGTCTGTTTCTCGGATGGATACCCATTAACATCT 1440
QY 1441 TGGACCTGGAAGAGTGTTCAGACAAAGTCTCCAACTCCACAGAGAGATCACAGAGGA 1500
DB 1441 TGGACCTGGAAGAGTGTTCAGACAAAGTCTCCAACTCCACAGAGAGATCACAGAGGA 1500
QY 1501 GTCTGGAATAGAAAGGCTTAAAGAAAGTGTGAGACAGTGGGTGTCAGCAGTACTCTA 1560
DB 1501 GTCTGGAATAGAAAGGCTTAAAGAAAGTGTGAGACAGTGGGTGTCAGCAGTACTCTA 1560
QY 1561 AACATGATGGAAGCCATAAAAGGTTCTGCTGCAAGTCTGTGCATACAAATTCCTTTGGC 1620
DB 1561 AACATGATGGAAGCCATAAAAGGTTCTGCTGCAAGTCTGTGCATACAAATTCCTTTGGC 1620
QY 1621 ACATCTTGTGAGACGATCCTTTTAACTCTCCAGGCCCCCTTCCCTTTTCATCCAGACAA 1680

DB 1621 ACATCTTGTGAGACGATCCTTTTAACTCTCCAGGCCCCCTTCCCTTTTCATCCAGACAA 1680
QY 1681 ATCTCATTTATGCAACAATTTGGTGTGTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1740
DB 1681 ATCTCATTTATGCAACAATTTGGTGTGTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1740
QY 1741 ATTTGTCAAGTAGTCAAAAGCAATTTAGGTATGAAAGCCAGCTACAGATGTCAGGGTG 1800
DB 1741 ATTTGTCAAGTAGTCAAAAGCAATTTAGGTATGAAAGCCAGCTACAGATGTCAGGGTG 1800
QY 1801 ACCGCTCTCTCAGATAATAGTACTCTTACGTTGATTTTTCAGAGATATGAATATCATCTC 1860
DB 1801 ACCGCTCTCTCAGATAATAGTACTCTTACGTTGATTTTTCAGAGATATGAATATCATCTC 1860
QY 1861 AAATGGGAGTTTCCAGAGAAAAATTTAGAGTTTGGAAAGTACTTAGGATCAGGTCTTTT 1920
DB 1861 AAATGGGAGTTTCCAGAGAAAAATTTAGAGTTTGGAAAGTACTTAGGATCAGGTCTTTT 1920
QY 1921 GGAAAGTGTATGAACGCAACAGCTTTATGGAATTAGCAAAACAGGAGTCTCAATCCAGGTT 1980
DB 1921 GGAAAGTGTATGAACGCAACAGCTTTATGGAATTAGCAAAACAGGAGTCTCAATCCAGGTT 1980
QY 1981 GCCCTCAAAATGCTGAAAGAAAAAGCAGACAGCTCTGAAAGAGAGGSCACTCATGTCAAG 2040
DB 1981 GCCCTCAAAATGCTGAAAGAAAAAGCAGACAGCTCTGAAAGAGAGGSCACTCATGTCAAG 2040
QY 2041 CTCAGATGATGACCCAGCTGGGAAGCCACAGAAATTTGTGAACCTCTGCTGGGGGCGTGC 2100
DB 2041 CTCAGATGATGACCCAGCTGGGAAGCCACAGAAATTTGTGAACCTCTGCTGGGGGCGTGC 2100
QY 2101 ACCTGTCCAGACCAATTTACTTCTGATTTTGAATTTGCTATGCTGATCTTCTCAAC 2160
DB 2101 ACCTGTCCAGACCAATTTACTTCTGATTTTGAATTTGCTATGCTGATCTTCTCAAC 2160
QY 2161 TATCTAAGAGTAAAGAGAAAAATTTTCAAGGACTTGGACAGAGATTTTCAAGGAAACAC 2220
DB 2161 TATCTAAGAGTAAAGAGAAAAATTTTCAAGGACTTGGACAGAGATTTTCAAGGAAACAC 2220
QY 2221 AATTTCAAGTTTATCCCACTTTTCCAATCACAATCCAAATTCAGCATCCCTGGTTCAGA 2280
DB 2221 AATTTCAAGTTTATCCCACTTTTCCAATCACAATCCAAATTCAGCATCCCTGGTTCAGA 2280
QY 2281 GAAGTTTCAGATACACCGGACTCGGATCAATCTCAGGCTTCATGGGAATTCATTTAC 2340
DB 2281 GAAGTTTCAGATACACCGGACTCGGATCAATCTCAGGCTTCATGGGAATTCATTTAC 2340
QY 2341 TCTCAAGATGAATTTGAATATGAACCAAAAGGCTGGAAGAGAGAGAGACTTGAAT 2400
DB 2341 TCTCAAGATGAATTTGAATATGAACCAAAAGGCTGGAAGAGAGAGAGACTTGAAT 2400
QY 2401 GTGCTTACATTTGAAGATCTTCTTGTGTTTGCATATCAAGTTTGCAGAGTGAATTT 2460
DB 2401 GTGCTTACATTTGAAGATCTTCTTGTGTTTGCATATCAAGTTTGCAGAGTGAATTT 2460
QY 2461 CTGGAATTTAAGTCTGTGTTTCAGAGACTGCGCGCCAGAGAGCTGCTGTCAACAC 2520
DB 2461 CTGGAATTTAAGTCTGTGTTTCAGAGAGCTGCGCGCCAGAGAGCTGCTGTCAACAC 2520
QY 2521 GGGAAAGTGGTGAAGATATGAGCTTTGGATTTGGCTCGAGATATCATGAGTGAATTTCCAC 2580
DB 2521 GGGAAAGTGGTGAAGATATGAGCTTTGGATTTGGCTCGAGATATCATGAGTGAATTTCCAC 2580
QY 2581 TATGTTGTGAGGGCAATGCCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2640
DB 2581 TATGTTGTGAGGGCAATGCCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2640
QY 2641 GAAGGCAATACACCAATTAAGAGTGTCTGGTTCATATGAATATTTACTGTGGGAAATC 2700
DB 2641 GAAGGCAATACACCAATTAAGAGTGTCTGGTTCATATGAATATTTACTGTGGGAAATC 2700
QY 2701 TTCTCCTTGTGAGTGAATCCTTACCTGGGCAATTCGGTGTGATGCTTAACCTTCTACAACTG 2760

2701	Db	TTCTCACTTGGGTGTAATCCCTTACCTGGCAATCCCGGTGATGCTAACTTCTACAAACTG	2767
2761	Qy	ATTCAAAATGGATTTAAAATGGATCAGCCATTTTATGCTACAGAAGAAATATACATTATA	2820
2761	Db	ATTCAAAATGGATTTAAAATGGATCAGCCATTTATGCTACAGAAGAAATATACATTATA	2820
2821	Qy	ATGCAATCTGCTGGGCTTTTGACTCAAGGAAACGGCCATCTTCCCTAATTTGACTTCG	2880
2821	Db	ATGCAATCTGCTGGGCTTTTGACTCAAGGAAACGGCCATCTTCCCTAATTTGACTTCG	2880
2881	Qy	TTTTTTAGGATGTCAGCTGGCAGATGCAGAGAGACGGATGTATCAGATGTGGGATGGCCGT	2940
2881	Db	TTTTTTAGGATGT CAGCTGGCAGATGCAGAGAGACGGATGTATCAGAAATGTGGATGGCCGT	2940
2941	Qy	GTTTCCGAATGTCTCTCACACCTACCAAAACAGGCGACCTTTTCAGCAGAGATGGATTTCG	3000
2941	Db	GTTTCCGAATGTCTCTCACACTACCAAAACAGGGACCTTTTCAGCAGAGATGGATTTCG	3000
3001	Qy	GGGCTACTCTCTCCGAGGCTCAGGTCGAAGATTCTGTAGAGGAACAATTTAGTTTAAAGG	3060
3001	Db	GGGCTACTCTCTCCGAGGCTCAGGTCGAAGATTCTGTAGAGGAACAATTTAGTTTAAAGG	3060
3061	Qy	ACTTCATCCCTCCACCTATCCCTTAACAGGCTGTAGATTACCAAAACAAGATTAAATTTTCAT	3120
3061	Db	ACTTCATCCCTCCACCTATCCCTTAACAGGCTGTAGATTACCAAAACAAGATTAAATTTTCAT	3120
3121	Qy	CACATAAAGAAAATCTATTATCAACTCTGCTTCACACAGACTTTTCTCTAGAGCCGCTCT	3180
3121	Db	CACATAAAGAAAATCTATTATCAACTCTGCTTCACACAGACTTTTCTCTAGAGCCGCTCT	3180
3181	Qy	CGCTTTACTCTTGTTTTCAAAGGACCTTTTGTAATAATCAAATCATCTGTCTCAGAGCCAG	3240
3181	Db	CGCTTTACTCTTGTTTTCAAAGGACCTTTTGTAATAATCAAATCATCTGTCTCAGAGCCAG	3240
3241	Qy	GAGAGCTGATAATGAACCTTTATTTGAGACATGTGATCTGCATCCAAAGCCCTTCTCAGGCCG	3300
3241	Db	GAGAGCTGATAATGAACCTTTATTTGAGACATGTGATCTGCATCCAAAGCCCTTCTCAGGCCG	3300
3301	Qy	GCCTTGATGTAATCTGTACCTCGAAGTACGATATATCTTGTAATAATCAATAAACAAGC	3360
3301	Db	GCCTTGATGTAATCTGTACCTCGAAGTACGATATATCTTGTAATAATCAATAAACAAGC	3360
3361	Qy	ATTTTGTCAAGGAGAAGCTAATATGATTTTTTTTAAAGTCTATGTTTTTAAAAATAATGTAAA	3420
3361	Db	ATTTTGTCAAGGAGAAGCTAATATGATTTTTTTTAAAGTCTATGTTTTTAAAAATAATGTAAA	3420
3421	Qy	TTTTTCAGCTATTTAGTGATATATTTTATGGGTGGGAATAAAATTTCTTACTACAG	3475
3421	Db	TTTTTCAGCTATTTAGTGATATATTTTATGGGTGGGAATAAAATTTCTTACTACAG	3475

```

1 RESULT 10
2 PCT-US95-03718-3
3 ; Sequence 3, Application PC/TUS9503718
4 ; GENERAL INFORMATION:
5 ;
6 ; APPLICANT: GENENTECH, INC.
7 ;
8 ; TITLE OF INVENTION: AGONIST ANTIBODIES AGAINST THE FLK2/FLT3 RECEPTOR AND USES THEREOF
9 ;
10 ; NUMBER OF SEQUENCES: 4
11 ;
12 ; CORRESPONDENCE ADDRESS:
13 ; ADDRESSEE: Genentech, Inc.
14 ; STREET: 460 Point San Bruno Blvd
15 ; CITY: South San Francisco
16 ; STATE: California
17 ; COUNTRY: USA
18 ; ZIP: 94080
19 ;
20 ; COMPUTER READABLE FORM:
21 ; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
22 ; COMPUTER: IBM PC compatible
23 ; OPERATING SYSTEM: PC-DOS/MS-DOS
24 ; SOFTWARE: patin (Genentech)
25 ;
26 ; CURRENT APPLICATION DATA:
27 ; APPLICATION NUMBER: PCT/US95/03718
28 ; FILING DATE:
29 ;

```

|||||
721 CATGAATATTTGGGACGGACATTAAGGTGCTGTGCCGAATGAATGACCTGGGCAGGGAATGC 780
Db
781 ACCAGGCTGTTTCAATATAGATCTTAATCAAACTCCTCAGACCAATTTGGCCACAATATATTT 840
Qy
781 ACCAGGCTGTTTCAATATAGATCTTAATCAAACTCCTCAGACCAATTTGGCCACAATATATTT 840
Db
841 CTTAAAGTGGGGAACCTTATGATTAAGGTGCAAGCTGTTTATGTAACCATGGATTC 900
Qy
841 CTTAAAGTGGGGAACCTTATGATTAAGGTGCAAGCTGTTTATGTAACCATGGATTC 900
Db
901 GGGCTCACCTGGGAATTTAGAAAAACAAGCACTCGAGGAGGGCAACTACTTTTGAGATGAGT 960
Qy
901 GGGCTCACCTGGGAATTTAGAAAAACAAGCACTCGAGGAGGGCAACTACTTTTGAGATGAGT 960
Db
961 ACCTATTCAACAAACAGAACTATGATACGGATTCGTTTGTGTAATCATAGTGSCA 1020
Qy
961 ACCTATTCAACAAACAGAACTATGATACGGATTCGTTTGTGTAATCATAGTGSCA 1020
Db
1021 AGAAACGACACCGGATACACTTGTTCCTCTTCAAAAGCATCCAGTCAATCAGCTTTG 1080
Qy
1021 AGAAACGACACCGGATACACTTGTTCCTCTTCAAAAGCATCCAGTCAATCAGCTTTG 1080
Db
1081 GTTACCATCGTAGGAAGGGATTTATAAATGCTACCAATTCAGGTGAAGATTTATGAAAT 1140
Qy
1081 GTTACCATCGTAGGAAGGGATTTATAAATGCTACCAATTCAGGTGAAGATTTATGAAAT 1140
Db
1141 GACCAATATGAAGAGTTTGTGTTTCTGTCAAGTTTAAAGGCTACCCACAATCAGATGT 1200
Qy
1141 GACCAATATGAAGAGTTTGTGTTTCTGTCAAGTTTAAAGGCTACCCACAATCAGATGT 1200
Db
1201 ACGTGGACCTTCTCGAAAATCAATTCCTGTGAGCAAAAAGGGCTTTGATTAACGGATAC 1260
Qy
1201 ACGTGGACCTTCTCGAAAATCAATTCCTGTGAGCAAAAAGGGCTTTGATTAACGGATAC 1260
Db
1261 AGCATATCCAAGTTTGAATCAATAGCACCAGCCAGGAGATATATATTCATGCAGAA 1320
Qy
1261 AGCATATCCAAGTTTGAATCAATAGCACCAGCCAGGAGATATATATTCATGCAGAA 1320
Db
1321 AATGATGATGCCAATTTACCAAAATGTTTCAGCTGAAATTAAGAAAGGAACCTCAAGTG 1380
Qy
1321 AATGATGATGCCAATTTACCAAAATGTTTCAGCTGAAATTAAGAAAGGAACCTCAAGTG 1380
Db
1381 CTCGAGAAGCATCGGCAAGTCAGGCGTCTGTTTCTCGATGGATACCCATACCATCT 1440
Qy
1381 CTCGAGAAGCATCGGCAAGTCAGGCGTCTGTTTCTCGATGGATACCCATACCATCT 1440
Db
1441 TGGACCTGGAAGAAGTGTTCAGACAAGTCTCCCACTGCAAGAGAGATCACAGAAGA 1500
Qy
1441 TGGACCTGGAAGAAGTGTTCAGACAAGTCTCCCACTGCAAGAGAGATCACAGAAGA 1500
Db
1501 GTCTGGAATAGAAAGGCTTAAAGAAAGTGTGAGCAGTGGGTGTCGAGCAGTACTCTA 1560
Qy
1501 GTCTGGAATAGAAAGGCTTAAAGAAAGTGTGAGCAGTGGGTGTCGAGCAGTACTCTA 1560
Db
1561 AACATGATGAGCAATTAAGGGTTCCTGGTCAAGTGTGTCATACAAATTCCTTGGC 1620
Qy
1561 AACATGATGAGCAATTAAGGGTTCCTGGTCAAGTGTGTCATACAAATTCCTTGGC 1620
Db
1621 ACATCTGTGAGCAATTCCTTTTAACTCTCAGGCGCTTCCCTTTTCAATCCAGACAAC 1680
Qy
1621 ACATCTGTGAGCAATTCCTTTTAACTCTCAGGCGCTTCCCTTTTCAATCCAGACAAC 1680
Db
1681 ATCTCATTTGCAACAAATTTGATGTTGCTCTCTTCAATGTCGTTTAAACCTGCTA 1740
Qy
1681 ATCTCATTTGCAACAAATTTGATGTTGCTCTCTTCAATGTCGTTTAAACCTGCTA 1740
Db
1741 ATTTGTCAAGATCAAAAAGCAATTTAGGTATGAAAGCCAGCTACAGATGGTACAGGTG 1800
Qy
1741 ATTTGTCAAGATCAAAAAGCAATTTAGGTATGAAAGCCAGCTACAGATGGTACAGGTG 1800
Db
1801 ACCGGCTCCTCAGATAATGAGTACTTCTACGTTGATTTTCAGAGAATATGAATATGATCTC 1860
Qy

1801 ACCGGCTCCTCAGATAATGAGTACTTCTACGTTGATTTTCAGAGAATATGAATATGATCTC 1860
Db
1861 AAATGGAGTTCCTCAAGAGAAATTTTAGAGTTTGGGAGGTACTAGGATCAGGTGCTTTT 1920
Qy
1861 AAATGGAGTTCCTCAAGAGAAATTTTAGAGTTTGGGAGGTACTAGGATCAGGTGCTTTT 1920
Db
1921 GGAAGAGTATGAACGCAACAGCTTATGGAATTTAGCAAAAACAGGAGTCTCAATCCAGGTT 1980
Qy
1921 GGAAGAGTATGAACGCAACAGCTTATGGAATTTAGCAAAAACAGGAGTCTCAATCCAGGTT 1980
Db
1981 GCCGTCAAAATGCTGAAGAAAAAGCAGACAGCTCTGAAAGAGAGGCACTCATGTTCAGAA 2040
Qy
1981 GCCGTCAAAATGCTGAAGAAAAAGCAGACAGCTCTGAAAGAGAGGCACTCATGTTCAGAA 2040
Db
2041 CTCAGAGTATGATGACCCAGCTGGGAGCCACAGAGATATTTGTGAACCTCTCGGGGCGGTGC 2100
Qy
2041 CTCAGAGTATGATGACCCAGCTGGGAGCCACAGAGATATTTGTGAACCTCTCGGGGCGGTGC 2100
Db
2101 ACCTGTCCAGGACCAATTTACTTGTATTTTGAATATCTGTTGCTATGTTGATCTTCTCAAC 2160
Qy
2101 ACCTGTCCAGGACCAATTTACTTGTATTTTGAATATCTGTTGCTATGTTGATCTTCTCAAC 2160
Db
2161 TATCTAAGAGTAAAGAGAAAAATTTTCAAGGACTTTGGACAGAGATTTTCAAGGAAACAC 2220
Qy
2161 TATCTAAGAGTAAAGAGAAAAATTTTCAAGGACTTTGGACAGAGATTTTCAAGGAAACAC 2220
Db
2221 AATTTTCAGTTTTTACCCCACTTTTCCAATCACAATCCAAATTCAGCATGCTGTTTCAAGA 2280
Qy
2221 AATTTTCAGTTTTTACCCCACTTTTCCAATCACAATCCAAATTCAGCATGCTGTTTCAAGA 2280
Db
2281 GAAGTTTCAGATACACCCGAGCTCGGATCAAAATCTCAGGGCTTCATGGGAATTCATTTTCA 2340
Qy
2281 GAAGTTTCAGATACACCCGAGCTCGGATCAAAATCTCAGGGCTTCATGGGAATTCATTTTCA 2340
Db
2341 TCTGAAGATGAATTTGAATATGAATAATGAATAATGAATAATGAATAATGAATAATGAATA 2400
Qy
2341 TCTGAAGATGAATTTGAATATGAATAATGAATAATGAATAATGAATAATGAATAATGAATA 2400
Db
2401 GTGCTTACATTTGAAGATCTTTCTTTGCTTGCATATCAAGTTTGCCTGCAAGAGATGAATTT 2460
Qy
2401 GTGCTTACATTTGAAGATCTTTCTTTGCTTGCATATCAAGTTTGCCTGCAAGAGATGAATTT 2460
Db
2461 CTGGAAATTTAAGTGTGTTTCAGAGACCTGGCCGCGCAGGAACGTCTGTGTCACCCAC 2520
Qy
2461 CTGGAAATTTAAGTGTGTTTCAGAGACCTGGCCGCGCAGGAACGTCTGTGTCACCCAC 2520
Db
2521 GGGAAAGTGTGAAGATATGTGACTTTTGGATTTGGCTCGAGATATCATGAGTGAATTTCCAC 2580
Qy
2521 GGGAAAGTGTGAAGATATGTGACTTTTGGATTTGGCTCGAGATATCATGAGTGAATTTCCAC 2580
Db
2581 TATGTTGTTCAGGGCAATGCGCTGCTGTGTAAATGGATGGCCCGCAAGAGCTGTTT 2640
Qy
2581 TATGTTGTTCAGGGCAATGCGCTGCTGTGTAAATGGATGGCCCGCAAGAGCTGTTT 2640
Db
2641 GAAGGCATCTACACATTAAGAGATGTCTGGTTCATATGGAATATTTACTGTTGGGAAATC 2700
Qy
2641 GAAGGCATCTACACATTAAGAGATGTCTGGTTCATATGGAATATTTACTGTTGGGAAATC 2700
Db
2701 TTCTCAGTTGTGTGAATCTTACCCTGCGATTCGGTTGATGCTAACTTCTTACAAACTG 2760
Qy
2701 TTCTCAGTTGTGTGAATCTTACCCTGCGATTCGGTTGATGCTAACTTCTTACAAACTG 2760
Db
2761 ATTCAAAATGGATTTAAATGGATCAGCCATTTTATGCTACAGAGAAATATACATTATA 2820
Qy
2761 ATTCAAAATGGATTTAAATGGATCAGCCATTTTATGCTACAGAGAAATATACATTATA 2820
Db
2821 ATGCAATCTGTGGGCTTTTGAATCTCAAGGAAACCGGCCATCTTCCCTTAATTTGACTTCG 2880
Qy
2821 ATGCAATCTGTGGGCTTTTGAATCTCAAGGAAACCGGCCATCTTCCCTTAATTTGACTTCG 2880
Db
2881 TTTTGTAGATCTCAGCTGGCAGATGTCAGAGAGCGATGCTATCAGAAATGTGATGSCCT 2940
Qy
2881 TTTTGTAGATCTCAGCTGGCAGATGTCAGAGAGCGATGCTATCAGAAATGTGATGSCCT 2940
Db

QY 2941 GTTTCGGAATGCTCCACACCTACCAAAACAGGCGACCTTTTCAGCAGAGAGATGATTTG 3000
DB 2941 GTTTCGGAATGCTCCACACCTACCAAAACAGGCGACCTTTTCAGCAGAGAGATGATTTG 3000
QY 3001 GGGCTACTCTCCGCGAGCTCAGTCGAGAGTCTGAGAGAACTTTAGTTTAAAGG 3060
DB 3001 GGGCTACTCTCCGCGAGCTCAGTCGAGAGTCTGAGAGAACTTTAGTTTAAAGG 3060
QY 3061 ACTTCATCCCTCCACCTATCCCTAACAGGCTGTAGATTACCAAAACAGATTTTTCAT 3120
DB 3061 ACTTCATCCCTCCACCTATCCCTAACAGGCTGTAGATTACCAAAACAGATTTTTCAT 3120
QY 3121 CACTAAAGAAATCTATTATCAACTGCTGCTTCCACAGACTTTTCTCAGAGCGCTCT 3180
DB 3121 CACTAAAGAAATCTATTATCAACTGCTGCTTCCACAGACTTTTCTCAGAGCGCTCT 3180
QY 3181 GCGTTTACTCTGTTTCAAGGGACTTTTGAAGTCAATCAATCAATCAATCAATCAAT 3240
DB 3181 GCGTTTACTCTGTTTCAAGGGACTTTTGAAGTCAATCAATCAATCAATCAATCAAT 3240
QY 3241 GAGGAGCTGATTAATCAACTTTTATGAGGACTTGTGATCCAGGCTTCTCAGGCGG 3300
DB 3241 GAGGAGCTGATTAATCAACTTTTATGAGGACTTGTGATCCAGGCTTCTCAGGCGG 3300
QY 3301 GCTTGAGTGAATTTGTTACCTGAGTACAGTATATTTTGTAAATACATAAAACAAAGC 3360
DB 3301 GCTTGAGTGAATTTGTTACCTGAGTACAGTATATTTTGTAAATACATAAAACAAAGC 3360
QY 3361 ATTTGCTAAGAGAGCTAATATGATTTTAACTGATGTTTAAATATATATGATAA 3420
DB 3361 ATTTGCTAAGAGAGCTAATATGATTTTAACTGATGTTTAAATATATATGATAA 3420
QY 3421 TTTTTCAGCTATTTAGTGTATATTTTATGCGTGGGAAATTTTCTACTACAG 3475
DB 3421 TTTTTCAGCTATTTAGTGTATATTTTATGCGTGGGAAATTTTCTACTACAG 3475

RESULT 11

US-08-183-211-1
; Sequence 1, Application US/08183211
; Patent No. 5618709
; GENERAL INFORMATION:
; APPLICANT: Alan M. Gewirtz, Donald Small, Curt I. Civin.
; TITLE OF INVENTION: ANTISENSE OLIGONUCLEOTIDES
; TITLE OF INVENTION: SPECIFIC FOR STK-1 AND METHOD FOR
; TITLE OF INVENTION: INHIBITING EXPRESSION OF THE STK-1 PROTEIN
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESS: SEIDEL GONDA LAVORGNA & MONACO
; STREET: Suite 1800, Penn Center Plaza
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: U.S.A.
; ZIP: 19102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/183,211
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Monaco, Daniel A.
; REGISTRATION NUMBER: 30,480
; REFERENCE/DOCKET NUMBER: 3957-15
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-8383

; TELEFAX: (215) 568-5549
; TELEX: No. 5618709e
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3476 nucleotides
; TYPE: nucleic acid
; STRANDEDNESS: single stranded
; TOPOLOGY: linear
; US-08-183-211-1

Query Match 99.2%; Score 3474.4; DB 1; Length 3476;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3475; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 CGAGCGCGCATCCGAGGCTGGCGCGGCCCTGGGGGACCCCGGCTCCGAGGCGCATG 60
DB 1 CGAGCGCGCATCCGAGGCTGGCGCGGCCCTGGGGGACCCCGGCTCCGAGGCGCATG 60
QY 61 CGGCGGTTGGCGCGGACGCGGCGCCGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
DB 61 CGGCGGTTGGCGCGGACGCGGCGCCGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
QY 121 TTTGGGACTATTACAAATCAAGATCTGCTGCTGATCAAGTGTGTTTAAATCAATCAATG 180
DB 121 TTTGGGACTATTACAAATCAAGATCTGCTGCTGATCAAGTGTGTTTAAATCAATCAATG 180
QY 181 AACAAATGATTCATAGTGGGGAAGTCATCATCATATCCCATGATFCAGAAATCCCGGAA 240
DB 181 AACAAATGATTCATAGTGGGGAAGTCATCATCATATCCCATGATFCAGAAATCCCGGAA 240
QY 241 GACCTCGGCTGCTGCTGAGACCCAGAGCTCAGGACAGTGTACGAAGTGGCGGTGTG 300
DB 241 GACCTCGGCTGCTGCTGAGACCCAGAGCTCAGGACAGTGTACGAAGTGGCGGTGTG 300
QY 301 GAAGTGGATGATCTGCTTCCATCACCTGCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCT 360
DB 301 GAAGTGGATGATCTGCTTCCATCACCTGCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCT 360
QY 361 TCCTGCTCTGGGCTCTTTAAGCACAGCTCCCTGAAATGCGCAGCCACATTTTGTATACAA 420
DB 361 TCCTGCTCTGGGCTCTTTAAGCACAGCTCCCTGAAATGCGCAGCCACATTTTGTATACAA 420
QY 421 AACAGAGGATGTTTCCATGCTCATTTTGAATAATGACAGAAACCCAGCTGGAGATAC 480
DB 421 AACAGAGGATGTTTCCATGCTCATTTTGAATAATGACAGAAACCCAGCTGGAGATAC 480
QY 481 CTACTTTTATTCAGAGTGAAGTACCAATTCACAAATATTTTACAGTACGTATAGA 540
DB 481 CTACTTTTATTCAGAGTGAAGTACCAATTCACAAATATTTTACAGTACGTATAGA 540
QY 541 AATACCTGCTTTTACACATTAAGAAAGACCTTACTTTAGAAAAATGAAAAACAGGACGCC 600
DB 541 AATACCTGCTTTTACACATTAAGAAAGACCTTACTTTAGAAAAATGAAAAACAGGACGCC 600
QY 601 CTGGTCTGCATATCTGAGAGCGTTCCAGAGCGATGCTGGAATGGGTGCTTTCGATTC 660
DB 601 CTGGTCTGCATATCTGAGAGCGTTCCAGAGCGATGCTGGAATGGGTGCTTTCGATTC 660
QY 661 CAGGGGGAAGCTGTAAGAGAAAGTCCAGCTGTTGTTTAAAGAGAGGAAAGTGCCTT 720
DB 661 CAGGGGGAAGCTGTAAGAGAAAGTCCAGCTGTTGTTTAAAGAGAGGAAAGTGCCTT 720
QY 721 CATGAATTTATTTGGGACGAGCATAGAGTGTGCTGCCAGAAATGAACTGGGCGAGGAATGC 780
DB 721 CATGAATTTATTTGGGACGAGCATAGAGTGTGCTGCCAGAAATGAACTGGGCGAGGAATGC 780
QY 781 ACCAGGCTGTTTCAATATAGATCTAAATCAAACTCCCTCAGACCAATTCGACCAATTTT 840
DB 781 ACCAGGCTGTTTCAATATAGATCTAAATCAAACTCCCTCAGACCAATTCGACCAATTTT 840
QY 841 CTTAAAGTAGGGGAAACCCCTTATGGATAAGTGTGCAAGCTGTTTCATGTGACCATGATTC 900
DB 841 CTTAAAGTAGGGGAAACCCCTTATGGATAAGTGTGCAAGCTGTTTCATGTGACCATGATTC 900


```

; Sequence 1, Application PC/TUS9500176A
; GENERAL INFORMATION:
; APPLICANT: Alan M. Gewirtz, Donald Small, Curt I. Givin.
; TITLE OF INVENTION: ANTISENSE OLIGONUCLEOTIDES
; TITLE OF INVENTION: SPECIFIC FOR SYK-1 AND METHOD FOR INHIBITING EXPRESSION OF THE
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEIDEL GONDA LAVORGNA & MONACO
; STREET: Suite 1800, Penn Center Plaza
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: U.S.A.
; ZIP: 19102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 720 KB
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/00176A
; FILING DATE: 6 January 1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/183,211
; FILING DATE: 14 January 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Monaco, Daniel A.
; REGISTRATION NUMBER: 30,480
; REFERENCE/DOCKET NUMBER: 3957-14 PC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-8383
; TELEFAX: (215) 568-5549
; TELEX: None
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3476 nucleotides
; TYPE: nucleic acid
; STRANDEDNESS: single stranded
; TOPOLOGY: linear
;
PCT-US95-00176A-1
Query Match 99.2%; Score 3474.4; DB 5; Length 3476;
Best Local Similarity 100.0%; Pred. No. 0;

```

```

Matches 3475; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 CGAGGGGCGCATCCGAGGGCTGGGCGGCGCCCTGGGGGAGACCCCGGGCTCGGAGGCCCATG 60
DB 1 CGAGGGGCGCATCCGAGGGCTGGGCGGCGCCCTGGGGGAGACCCCGGGCTCGGAGGCCCATG 60
QY 61 CGGCGGTTGGCGGCGGAGCGGCGGCGCCGCTGCTCGGTTGTTTCTGCAATGATA 120
DB 61 CGGCGGTTGGCGGCGGAGCGGCGGCGCCGCTGCTCGGTTGTTTCTGCAATGATA 120
QY 121 TTGGGACTATTACAATCAAGATCTGCTGCTGATCAAGTGTGTTTAAATCAATCAATG 180
DB 121 TTGGGACTATTACAATCAAGATCTGCTGCTGATCAAGTGTGTTTAAATCAATCAATG 180
QY 181 AACAAATGATTCATCAGTGGGAAAGTCATCATATCCATGATGATGATGATGATGATG 240
DB 181 AACAAATGATTCATCAGTGGGAAAGTCATCATATCCATGATGATGATGATGATGATG 240
QY 241 GACCTGGGTTGCGTTGAGACCCGAGCTCAGGGACAGTGTACGAGCTGCGCTGTG 300
DB 241 GACCTGGGTTGCGTTGAGACCCGAGCTCAGGGACAGTGTACGAGCTGCGCTGTG 300
QY 301 GAAGTGATGATGATGCTGCTTCCATCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTG 360
DB 301 GAAGTGATGATGATGCTGCTTCCATCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTG 360
QY 361 TCCTGCTCTGGGTTTAAAGCAGCTCCCTGAAATGCCAGCCACATTTTGAATTAACA 420
DB 361 TCCTGCTCTGGGTTTAAAGCAGCTCCCTGAAATGCCAGCCACATTTTGAATTAACA 420
QY 421 AACAGAGGAGTTGTTTCCATGCTGCTTGAATGACAGAAACCCAGCTGGAGAAATAC 480
DB 421 AACAGAGGAGTTGTTTCCATGCTGCTTGAATGACAGAAACCCAGCTGGAGAAATAC 480
QY 481 CTACTTTTATTCAGAGTGAAGTACCAATTTACACATATTTTACAGTGAAGTAAAGA 540
DB 481 CTACTTTTATTCAGAGTGAAGTACCAATTTACACATATTTTACAGTGAAGTAAAGA 540
QY 541 AATACCTGCTTTTACACATTAAGAGACCTTTACTTTAGAAAATGGAAGAACCCAG 600
DB 541 AATACCTGCTTTTACACATTAAGAGACCTTTACTTTAGAAAATGGAAGAACCCAG 600
QY 601 CTGCTGCTATCTGAGAGGCTTCCAGAGCGGATCGTGGAAATGGGTCCTTTGCAATCA 660
DB 601 CTGCTGCTATCTGAGAGGCTTCCAGAGCGGATCGTGGAAATGGGTCCTTTGCAATCA 660
QY 661 CAGGGGAAAGCTGTAAAGAGAAAGTCCAGCTGTTGTTTAAAGAGGAGAAAGTGTCT 720
DB 661 CAGGGGAAAGCTGTAAAGAGAAAGTCCAGCTGTTGTTTAAAGAGGAGAAAGTGTCT 720
QY 721 CATGAATTTATTTGGGACGGACATTAAGGTGCTGCTGCGAGAAATGAATCGGCGAG 780
DB 721 CATGAATTTATTTGGGACGGACATTAAGGTGCTGCTGCGAGAAATGAATCGGCGAG 780
QY 781 ACCAGGCTGTTTCAATAGATCTAAATCAAACTCCTCCTCAGACCAATTCGCCCAAT 840
DB 781 ACCAGGCTGTTTCAATAGATCTAAATCAAACTCCTCCTCAGACCAATTCGCCCAAT 840
QY 841 CTTAAAGTAGGGGAAACCTTTATGATGATGATGATGATGATGATGATGATGATGATG 900
DB 841 CTTAAAGTAGGGGAAACCTTTATGATGATGATGATGATGATGATGATGATGATGATG 900
QY 901 GGGCTCACCTGGGAAATTAGAAAACAAAGCACTCGAGGAGGGCAACTACTTTGAGATG 960
DB 901 GGGCTCACCTGGGAAATTAGAAAACAAAGCACTCGAGGAGGGCAACTACTTTGAGATG 960
QY 961 ACCTATTCAACAAACAGAACTATGATGATGATGATGATGATGATGATGATGATGATG 1020
DB 961 ACCTATTCAACAAACAGAACTATGATGATGATGATGATGATGATGATGATGATGATG 1020
QY 1021 AGAAACGACCGGATACACTACATTTGCTTCAAGAGCATCCAGCTCAATCAGCTTTG 1080
DB 1021 AGAAACGACCGGATACACTACATTTGCTTCAAGAGCATCCAGCTCAATCAGCTTTG 1080

```


Db 3241 GAGGAGCTGATAATGAACCTTTATTGGAGCAATTGATCTGCATCCAAAGGCCCTTCTCAGGCGG 3300
Qy 3301 GCTTCAGTGAATTTGTGCTCAAGTACAGTAGTATTTCTTTGTAATAATACATAAAACAAAAGC 3360
Db 3301 GCTTCAGTGAATTTGTGCTCAAGTACAGTAGTATTTCTTTGTAATAATACATAAAACAAAAGC 3360
Qy 3361 ATTTTGTAAAGGAGAAGCTAAATATGATTTTTTTAAAGTCTAAGTTTTTTAAATAATATGTAAA 3420
Db 3361 ATTTTGTAAAGGAGAAGCTAAATATGATTTTTTTAAAGTCTAAGTTTTTTAAATAATATGTAAA 3420
Qy 3421 TTTTTCAGCTATTAGTAGTATATTTTATGGTGGGAATAAAATTTCTACTACAGA 3476
Db 3421 TTTTTCAGCTATTAGTAGTATATTTTATGGTGGGAATAAAATTTCTACTACAGA 3476

RESULT 13
US-08-222-616-22
; Sequence 22, Application US/08222616
; Patent No. 5635177
; GENERAL INFORMATION:
; APPLICANT: Bennett, Brian D.
; APPLICANT: Goeddel, David
; APPLICANT: Lee, James M.
; APPLICANT: Matthews, William
; APPLICANT: Tsai, Siao Ping
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: PROTEIN TYROSINE KINASE AGONIST
; NUMBER OF INVENTION: ANTIBODIES
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/222,616
; FILING DATE: 4-APR-1994
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/00586
; FILING DATE: 22-JAN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/826935
; FILING DATE: 22-JAN-1992
; NAME: Lee, Wendy M.
; ATTORNEY/AGENT INFORMATION:
; REGISTRATION NUMBER:
; REFERENCE/DOCKET NUMBER: 821P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-1994
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3120 bases
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-222-616-22
Query Match 88.1%; Score 3083.2; DB 1; Length 3120;
Best/Local Similarity 99.3%; Pred. No. 0;
Matches 3097; Conservative 0; Mismatches 23; Indels 0; Gaps 0;
Qy 53 ATCCGCGCTTGGCGCGACGGGCGACCGTGCCTGCTGTTGTTTCTGCAATG 117

Db 1 ATGAGAGGTTGGCGCGGACGGCGGCGCAGCTGCGCGTCTGTTGTTTCTGCAATG 60
Qy 118 ATATTTGGGACTATTACAAATCAAGATCTGCTGTGATCAAGTGTGTTTTTAATCAATCAT 177
Db 61 ATATTTGGGACTATTACAAATCAAGATCTGCTGTGATCAAGTGTGTTTTTAATCAATCAT 120
Qy 178 AAGAACATGATTCATCAGTGGGGAAGTCAATCATCATATATCCCATGTTATCAGAAATCCCCG 237
Db 121 AAGAACATGATTCATCAGTGGGGAAGTCAATCATCATATATCCCATGTTATCAGAAATCCCCG 180
Qy 238 GAAGACCTCGGCTGTGCGTTGAGACCCAGAGCTCAGGACAGCTGTACGAAGCTGCGCGCT 297
Db 181 GAAGACCTCGGCTGTGCGTTGAGACCCAGAGCTCAGGACAGCTGTACGAAGCTGCGCGCT 240
Qy 298 GTGGAAGTGGATGTATCTGCTTCCATCACACTGCAAGTGTGTTGATGTTTGAATTTA 357
Db 241 GTGGAAGTGGATGTATCTGCTTCCATCACACTGCAAGTGTGTTGATGTTTGAATTTA 300
Qy 358 ATTTCTGCTCTGCGTCTTTAAGACAGCTCCCTGAAATGGCAGCCACATTTTGAATTTA 417
Db 301 ATTTCTGCTCTGCGTCTTTAAGACAGCTCCCTGAAATGGCAGCCACATTTTGAATTTA 360
Qy 418 CAAACAGAGGAGTGTGTTTCCATGTTGTTTGAATAATGACAGAAACCCCAAGCTGGAGAA 477
Db 361 CAAACAGAGGAGTGTGTTTCCATGTTGTTTGAATAATGACAGAAACCCCAAGCTGGAGAA 420
Qy 478 TACTACTTTTTTATTCAGAGTGAAGCTACCAATTAACAATATTTGTTTACAGTGAGTATA 537
Db 421 TACTACTTTTTTATTCAGAGTGAAGCTACCAATTAACAATATTTGTTTACAGTGAGTATA 480
Qy 538 AGAATACCTCTGTTTACATTAAGAAGACCTTACTTTAGAAAATGGAACACAGGAC 597
Db 481 AGAATACCTCTGTTTACATTAAGAAGACCTTACTTTAGAAAATGGAACACAGGAC 540
Qy 598 GCGCTGCTGCATATCTCAGAGCGTTCCAGAGCCGATCGTGAATGGGTGCTTTGCGAT 657
Db 541 GCGCTGCTGCATATCTCAGAGCGTTCCAGAGCGGATCTCGAATGGGTGCTTTGCGAT 600
Qy 658 TCACAGGGGAAAGCTGTAAAGAAAGTCCAGCTGTTGTTTAAAGAGGAGGAAAAAGTG 717
Db 601 TCACAGGGGAAAGCTGTAAAGAAAGTCCAGCTGTTGTTTAAAGAGGAGGAAAAAGTG 660
Qy 718 CTTTCAATGAATTTTGGGACGACATAAGTGTCTGTGCCAGAAATGAATCGGCGAGGAA 777
Db 661 CTTTCAATGAATTTTGGGACGACATAAGTGTCTGTGCCAGAAATGAATCGGCGAGGAA 720
Qy 778 TGCACAGGCTGTTTCAATAGATCTAAATCAAACTCCTCAGACCAATTTGCCACAATTA 837
Db 721 TGCACAGGCTGTTTCAATAGATCTAAATCAAACTCCTCAGACCAATTTGCCACAATTA 780
Qy 838 TTTCTTTAAAGTAGGGGAAACCTTTATGATAAGGTGCAAGCTGTTTCAATGTGAACCATGGA 897
Db 781 TTTCTTTAAAGTAGGGGAAACCTTTATGATAAGGTGCAAGCTGTTTCAATGTGAACCATGGA 840
Qy 898 TTCGGGCTCACCTGGGAATTTAGAAAAACAAAGCTCGAGAGGGGCAACTCTTTGAGATG 957
Db 841 TTCGGGCTCACCTGGGAATTTAGAAAAACAAAGCTCGAGAGGGGCAACTCTTTGAGATG 900
Qy 958 AGTACCTTTTCAACAAACAGAACTATGATACGGATTTCTGTTGCTTTTGTATCATCAGTG 1017
Db 901 AGTACCTTTTCAACAAACAGAACTATGATACGGATTTCTGTTGCTTTTGTATCATCAGTG 960
Qy 1018 GCAAGAAACGACACCGGATACTACCTTGTCTCTTCAAGAGATCCAGTCAATCAGCT 1077
Db 961 GCAAGAAACGACACCGGATACTACCTTGTCTCTTCAAGAGATCCAGTCAATCAGCT 1020
Qy 1078 TTGGTTACCATCGTAGGAAAGGATTTATAAATGCTACCAATTCAGTGAGGATTTATGAA 1137
Db 1021 TTGGTTACCATCGTAGGAAAGGATTTATAAATGCTACCAATTCAGTGAGGATTTATGAA 1080
Qy 1138 ATTGACCAATATCAAGAGTGTGTTTCTGTCAGGTTTAAAGCTACCCCAATCAGA 1197

Db 1081 ATTGACCAATATGAAGAGTTTGTGTTCTGTGAGGTTTAAAGCCTACCCACAAATCAGA 1140
Qy 1198 TGTACGTGACCTTCTCTCGAAATCAATTTCTGTGTGAGCAAAAGGCTTGTATACGGA 1257
Db 1141 TGTACGTGACCTTCTCTCGAAATCAATTTCTGTGTGAGCAAAAGGCTTGTATACGGA 1200
Qy 1258 TACAGCATATCCAAATTTGCAATCATAGCAACCCAGCAGGAGAAATATATATTTCCATGCA 1317
Db 1201 TACAGCATATCCAAATTTGCAATCATAGCAACCCAGCAGGAGAAATATATATTTCCATGCA 1260
Qy 1318 GAAATGATGATGCCCAATTTACCAAAATGTTTCAAGCTGAAATATAGAGGAAACCTCAA 1377
Db 1261 GAAATGATGATGCCCAATTTACCAAAATGTTTCAAGCTGAAATATAGAGGAAACCTCAA 1320
Qy 1378 GTGCTCGCAGAGCATCGGCAAGTCTGAGCAAGTCTCCCACTGCACAGAGAGATCACAGAA 1437
Db 1321 GTGCTCGCAGAGCATCGGCAAGTCTGAGCAAGTCTCCCACTGCACAGAGAGATCACAGAA 1380
Qy 1438 TCTTGGACCTGGAAGAGTGTTCAGACAACTCTCCCACTGCACAGAGAGATCACAGAA 1497
Db 1381 TCTTGGACCTGGAAGAGTGTTCAGACAACTCTCCCACTGCACAGAGAGATCACAGAA 1440
Qy 1498 GGAGTCTGGAATAGAAAGGCTAAACAGAAAGTGTTCGACAGTGGGTGTCGACGAGTACT 1557
Db 1441 GGAGTCTGGAATAGAAAGGCTAAACAGAAAGTGTTCGACAGTGGGTGTCGACGAGTACT 1500
Qy 1558 CTAACATGAGTGAAGCCATAAAGGGTCTCGTCAAGTGTGTCGATACAAATTCCTT 1617
Db 1501 CTAACATGAGTGAAGCCATAAAGGGTCTCGTCAAGTGTGTCGATACAAATTCCTT 1560
Qy 1618 GGCACATCTGTGAGACGATCTTTTAACTCTCCAGGCCCTTCCCTTTCATCCAGAC 1677
Db 1561 GGCACATCTGTGAGACGATCTTTTAACTCTCCAGGCCCTTCCCTTTCATCCAGAC 1620
Qy 1678 AACATCTCATCTTATGCAACAAATGTGTGTGTCTCTCTCTTCAATGTGCTTTTAAACCTG 1737
Db 1621 AACATCTCATCTTATGCAACAAATGTGTGTGTCTCTCTCTTCAATGTGCTTTTAAACCTG 1680
Qy 1738 CTAATTTGTACAGTACAAAAGCAATTTAGGTATGAAGCCAGCTACAGATGGTACAG 1797
Db 1681 CTAATTTGTACAGTACAAAAGCAATTTAGGTATGAAGCCAGCTACAGATGGTACAG 1740
Qy 1798 GTGACCGGCTCTCAGATATGAGTACTTCTAGTGTGATTTTTCAGAGAAATGATATGAT 1857
Db 1741 GTGACCGGCTCTCAGATATGAGTACTTCTAGTGTGATTTTTCAGAGAAATGATATGAT 1800
Qy 1858 CTCAATGGGAGTTTCCAGAGAGAAATTTAGAGTTTGGGAAGTACTAGGATCAGGTGCT 1917
Db 1801 GTCAATGGGAGTTTCCAGAGAGAAATTTAGAGTTTGGGAAGTACTAGGATCAGGTGCT 1860
Qy 1918 TTTGGAAGAGTATGAAACGCAACAGCTTATGGAATTTAGCAAAACAGGAGTCTCAATCCAG 1977
Db 1861 TTTGGAAGAGTATGAAACGCAACAGCTTATGGAATTTAGCAAAACAGGAGTCTCAATCCAG 1920
Qy 1978 GTTGGCGTCAAAATGCTGAAAGAAAGCAGACAGCTCTGAAAGAGAGGACATCATGTCA 2037
Db 1921 GTTACCGTCAAAATGCTGAAAGAAAGCAGACAGCTCTGAAAGAGAGGACATCATGTCA 1980
Qy 2038 GAACTCAAGATGATGACCCAGCTGGGAAGCCAGCAATATTTGTGAACCTGCTGGGGCG 2097
Db 1981 GAACTCAAGATGATGACCCAGCTGGGAAGCCAGCAATATTTGTGAACCTGCTGGGGCG 2040
Qy 2098 TGCACTGTGAGGACCAATTTACTTGTATTTTGAATCTGTTGCTATGTTGATCTTCTC 2157
Db 2041 TGCACTGTGAGGACCAATTTACTTGTATTTTGAATCTGTTGCTATGTTGATCTTCTC 2100
Qy 2158 AACTATCTAAGAGTAAAGAGAGAAATTTTCAGGACTTGGACAGAGATTTTTCAGGAA 2217
Db 2101 AACTATCTAAGAGTAAAGAGAGAAATTTTCAGGACTTGGACAGAGATTTTTCAGGAA 2160
Qy 2218 CACAATTTTCAAGTTTACCCCACTTTTCAATCATATCCAAATTTCCAGCATGCTGGTTCA 2277
Db 2161 CACAATTTTCAAGTTTACCCCACTTTTCAATCATATCCAAATTTCCAGCATGCTGGTTCA 2220

Qy 2278 AGAAGATTCAGATACACCCGAGCTCGGATCAATCTCAGGGCTTCATGGAAATCATTT 2337
Db 2221 AGAAGATTCAGATACACCCGAGCTCGGATCAATCTCAGGGCTTCATGGAAATCATTT 2280
Qy 2338 CACTCTGAAGATGAAATTTGAATATGAAAAACAAAAAGGCTGGAAGAGAGGAGACTTG 2397
Db 2281 CACTCTGAAGATGAAATTTGAATATGAAAAACAAAAAGGCTGGAAGAGAGGAGACTTG 2340
Qy 2398 AATGTGCTTACATTTGAAGATCTTCTTGTGTCATATCAAGTTGGCCAAAGAAATGAA 2457
Db 2341 AATGTGCTTACATTTGAAGATCTTCTTGTGTCATATCAAGTTGGCCAAAGAAATGAA 2400
Qy 2458 TTTCTGGAATTTAAGTCTGTGTTTCAAGAGACTGCGCCGAGGAACTGCTGTACCC 2517
Db 2401 TTTCTGGAATTTAAGTCTGTGTTTCAAGAGACTGCGCCGAGGAACTGCTGTACCC 2460
Qy 2518 CACGGGAAAGTGTGAAGATATGTGACTTTGGATTGGCTCGAGATATCATGAGTGTATCC 2577
Db 2461 CACGGGAAAGTGTGAAGATATGTGACTTTGGATTGGCTCGAGATATCATGAGTGTATCC 2520
Qy 2578 AACTATTTGTCTAGGGGCAATGCCGCTCTGCTGTAAAAATGGATGGCCCGGAAAGCCTG 2637
Db 2521 AACTATTTGTCTAGGGGCAATGCCGCTCTGCTGTAAAAATGGATGGCCCGGAAAGCCTG 2580
Qy 2638 TTTGAAAGGCAATCTACACCATTAAGAGTGTCTGTGTCATATGGAATATTTACTGTGGAA 2697
Db 2581 TTTGAAAGGCAATCTACACCATTAAGAGTGTCTGTGTCATATGGAATATTTACTGTGGAA 2640
Qy 2698 ATCTTCTCACTTGGTGTGAATCTTACCCCTGGCAATCCCGTGTGATGCTTCTACAAA 2757
Db 2641 ATCTTCTCACTTGGTGTGAATCTTACCCCTGGCAATCCCGTGTGATGCTTCTACAAA 2700
Qy 2758 CTGATTTCAAAATGGAATTTAAATGGATCAGCCATTTTATGCTACAGAGAAATATACAT 2817
Db 2701 CTGATTTCAAAATGGAATTTAAATGGATCAGCCATTTTATGCTACAGAGAAATATACAT 2760
Qy 2818 ATATGCAATCTCTGCTGGGCTTTTGAATCAAGGAAACGGCCATCTTCCCTAAATTTGACT 2877
Db 2761 ATATGCAATCTCTGCTGGGCTTTTGAATCAAGGAAACGGCCATCTTCCCTAAATTTGACT 2820
Qy 2878 TCGTTTTTATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 2937
Db 2821 TCGTTTTTATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 2880
Qy 2938 CGTGTTCGGAATGTCTCTCAGCTTACCAAAACAGGAGCTTTTTCAGCAGAGAGATGAT 2997
Db 2881 CGTGTTCGGAATGTCTCTCAGCTTACCAAAACAGGAGCTTTTTCAGCAGAGAGATGAT 2940
Qy 2998 TTGGGGCTACTCTCTCCGAGGCTCAGGTGGAAGATTTCTGAGAGGAAACAATTTAGTTT 3057
Db 2941 TTGGGGCTACTCTCTCCGAGGCTCAGGTGGAAGATTTCTGAGAGGAAACAATTTAGTTT 3000
Qy 3058 AGGACTTCACTCTCCCTCACTTCCCTTAAAGGCTGTAGATTTACAAACAGAGATTAATTT 3117
Db 3001 AGGACTTCACTCTCCCTCACTTCCCTTAAAGGCTGTAGATTTACAAACAGAGATTAATTT 3060
Qy 3118 CATCACTAAAGAAATCTTATTAATCACTGCTCTTCAAGAGCTTTTCTCTAGAGCGG 3177
Db 3061 CATCACTAAAGAAATCTTATTAATCACTGCTCTTCAAGAGCTTTTCTCTAGAGAGCG 3120

RESULT 14

US-08-446-648-22
; Sequence 22, Application US/08446648
; Patent No. 6331302
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Bennett, Brian D.
; APPLICANT: Goeddel, David
; APPLICANT: Lee, James M.
; APPLICANT: Matthews, William
; APPLICANT: Tsai, Siao Ping

Qy	1618	GGCACATCTTGTGAGACGATCCTCTTTTAAAC	CTCTCCAGGCCCTTCCCTTTTCATCCCAAGAC	1677
Db	1561	GGCACATCTTGTGAGACGATCCTCTTTTAAAC	CTCTCCAGGCCCTTCCCTTTTCATCCCAAGAC	1620
Qy	1678	AACATCTCATCTTATGCAACAA	TTTGGTGTGTCCTCTCTTCAATGTCGTTTTTAAACCCCTG	1737
Db	1621	AACATCTCATCTTATGCAACAA	TTTGGTGTGTCCTCTCTTCAATGTCGTTTTTAAACCCCTG	1680
Qy	1738	CTAATTTGTCACAAGTACAAAAAGCAAT	TTAGGTATGAAACCCAGCTTACAGATGGTACAG	1797
Db	1681	CTAATTTGTCACAAGTACAAAAAGCAAT	TTAGGTATGAAACCCAGCTTACAGATGGTACAG	1740
Qy	1798	GTGACCGGCTCCTCAGATAATGAGTAGTCTCTCA	CGTTGATTTACAGAGAAATAGTAATATGAT	1857
Db	1741	GTGACCGGATCCTCAGATAATGAGTAGTCTCTCA	CGTTGATTTACAGAGAAATAGTAATATGAT	1800
Qy	1858	CTCAATTTGGGAGTTTCCACAGAGAAA	TTTTAGAGTTTTGGGAAGGTACTAGGATCAGGTGCT	1917
Db	1801	GTCAATTTGGGAGTTTCCAGAGAGAAA	TTTTAGAGTTTTGGGAAGGTACTAGGATCAGGTGCT	1860
Qy	1918	TTTTGGAAAAGTGATGAACGCAACAGCTTTAT	TGGAATTTAGCAAAAAACAGAGGTCTCAATCCAG	1977
Db	1861	TTTTGGAAAAGTGATGAACGCAACAGCTTTAT	TGGAATTTAGCAAAAAACAGAGGTCTCAATCCAG	1920
Qy	1978	GTTCGCGCTCAAAATGCTGAAAGAAA	AAGACAGACTCTGAAAGAGAGGCACTCATGTCA	2037
Db	1921	GTTCACGCTCAAAATGCTGAAAGAAA	AAGACAGACTCTGAAAGAGAGGCACTCATGTCA	1980
Qy	2038	GAACTCNAGATGATGACCCAGCTGGGAGGCCA	CGAGNAATTTCTGAACCTGCTGGGGGGG	2097
Db	1981	GAACTCNAGATGATGACCCAGCTGGGAGGCCA	CGAGNAATTTCTGAACCTGCTGGGGGGG	2040
Qy	2098	TGCACACTGTCCAGGACCAATTTACTTGAT	TTTTTGAATCTGTTGCTATGSGTATCTTCTC	2157
Db	2041	TGCACACTGTCCAGGACCAATTTACTTGAT	TTTTTGAATCTGTTGCTATGSGTATCTTCTC	2100
Qy	2158	AACATCTAAGAAGTAAAGAGAAAAATTT	CACAGGACTTTGGAACAGAGATTTTCAAGGAA	2217
Db	2101	AACATCTAAGAAGTAAAGAGAAAAATTT	CACAGGACTTTGGAACAGAGATTTTCAAGGAA	2160
Qy	2218	CACAAATTTCAATTTTACCCCACTTTTCCAT	CAATCCAAATTTCCAGCATGCCCTGGTTCA	2277
Db	2161	CACAAATTTCAATTTTACCCCACTTTTCCAT	CAATCCAAATTTCCAGCATGCCCTGGTTCA	2220
Qy	2278	AGAGAGTTCCAGATACACCCGGACTCGGAT	CAAAATCTCAGGGCTTCATGGGAATTCATTT	2337
Db	2221	AGAGAGTTCCAGATACACCCGGACTCGGAT	CAAAATCTCAGGGCTTCATGGGAATTCATTT	2280
Qy	2338	CACCTCTGAAGATGAAATTTGAATATGAAAA	CCAAAAAGGCTGGAAGAGAGGAGGACTTG	2397
Db	2281	CACCTCTGAAGATGAAATTTGAATATGAAAA	CCAAAAAGGCTGGAAGAGAGGAGGACTTG	2340
Qy	2398	AATGTGCTTACATTTGAAGATCTCTTTGCT	TTTGCATATCAAGTTGCCAAGGAATGGAA	2457
Db	2341	AATGTGCTTACATTTGAAGATCTCTTTGCT	TTTGCATATCAAGTTGCCAAGGAATGGAA	2400
Qy	2458	TTTTCTGGAATTTAAGTCGTGTGTTCCAGAG	AGACTGCGCCAGAGGAAGCTGTTGTCACC	2517
Db	2401	TTTTCTGGAATTTAAGTCGTGTGTTCCAGAG	AGACTGCGCCAGAGGAAGCTGTTGTCACC	2460
Qy	2518	CACGGGAAAGTGGTGAAGATATGTGACT	TTTGGATTTGGCTCGAGATATCATGAGTGATTC	2577
Db	2461	CACGGGAAAGTGGTGAAGATATGTGACT	TTTGGATTTGGCTCGAGATATCATGAGTGATTC	2520
Qy	2578	AACATATGTTGTACGGGGCAATGCCCGT	CTGCTGTAAAAATGGAATGGCCCCCGAAGGCTG	2637
Db	2521	AACATATGTTGTACGGGGCAATGCCCGT	CTGCTGTAAAAATGGAATGGCCCCCGAAGGCTG	2580
Qy	2638	TTTTGAAGGCATCTACACCAATTAAGAGT	GTGTCGTGATATGGAATTTACTGTGGGAA	2697
Db	2581	TTTTGAAGGCATCTACACCAATTAAGAGT	GTGTCGTGATATGGAATTTACTGTGGGAA	2640
Qy	2698	ATCTTCTCACTTGGTGTGAAATCCTTTA	CCCTGGCAATTCGGGTGATGTAACTCTTACAAA	2757

[illegible]

RESULT 15
PCT-US95-04228-22
Sequence 22, Application PC/TUS9504228
GENERAL INFORMATION:
APPLICANT: Genentech, Inc.
APPLICANT: Bennett, Brian D.
APPLICANT: Goeddel, David
APPLICANT: Lee, James M.
APPLICANT: Matthews, William
APPLICANT: Tsai, Siao Ping
APPLICANT: Wood, William I.
TITLE OF INVENTION: PROTEIN TYROSINE KINASE AGONIST ANTIBODIES
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/04228
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/222616
FILING DATE: 04-APR-1994
ATTORNEY/AGENT INFORMATION:
NAME: Wendy M. Lee
REGISTRATION NUMBER: 00,000
REFERENCE/DOCKET NUMBER: 821P3PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-1994
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 22:

Search completed: August 28, 2003, 05:51:59
Job time : 293.45 secs

Qy	2098	TGCACACTGTCAGGACCAATTTACTTGTATTTTGAATACTTGTGCTATGCTGATCTTCTC	2157
Db	2041	TGCACACTGTCAGGACCAATTTACTTGTATTTTGAATACTTGTGCTATGCTGATCTTCTC	2100
Qy	2158	AACTATCTAAGAGTAAAGAGAGAAAAATTTACAGGACTTGGACAGAGATTTTCAAGGAA	2217
Db	2101	AACTATCTAAGAGTAAAGAGAGAAAAATTTACAGGACTTGGACAGAGATTTTCAAGGAA	2160
Qy	2218	CACAAATTTTCAGTTTACCCCACTTTCGAATCACATCCAAATTCAGCATGCCGTGTTCA	2277
Db	2161	CACAAATTTTCAGTTTACCCCACTTTCGAATCACATCCAAATTCAGCATGCCGTGTTCA	2220
Qy	2278	AGAGAAGTTTCAGATACACCCGCACTCGGATCAATCTCAGGGCTTCATGGGAATTCATTT	2337
Db	2221	AGAGAAGTTTCAGATACACCCGCACTCGGATCAATCTCAGGGCTTCATGGGAATTCATTT	2280
Qy	2338	CACTCTGAAGATGAATTAATATGAAAAACCAAAAAAGGCTGGAAGAGAGGAGACTTG	2397
Db	2281	CACTCTGAAGATGAATTAATATGAAAAACCAAAAAAGGCTGGAAGAGAGGAGACTTG	2340
Qy	2398	AATGTGCTTACATTTGAAGATCTTCTTGTGCTTGCATATCAAGTTGCCAAAGAAATGAA	2457
Db	2341	AATGTGCTTACATTTGAAGATCTTCTTGTGCTTGCATATCAAGTTGCCAAAGAAATGAA	2400
Qy	2458	TTTCTGGAAATTTAAGTCGTGCTTTCACAGAGACCTGGCCGCCAGGAACGTGCTTGCACC	2517
Db	2401	TTTCTGGAAATTTAAGTCGTGCTTTCACAGAGACCTGGCCGCCAGGAACGTGCTTGCACC	2460
Qy	2518	CACGGGAAGTGGTGAAGATATGTGACTTTTGGATTGGCTCGAGATATCATGAGTGATTC	2577
Db	2461	CACGGGAAGTGGTGAAGATATGTGACTTTTGGATTGGCTCGAGATATCATGAGTGATTC	2520
Qy	2578	AACTATGTTGTCAGGGGCAATGCCCGTCTGCTGTAAATGATGGCCGCCGAAAGCCCTG	2637
Db	2521	AACTATGTTGTCAGGGGCAATGCCCGTCTGCTGTAAATGATGGCCGCCGAAAGCCCTG	2580
Qy	2638	TTTGAAGGCATCTACACCATTAAGAGTGTCTGCTCATATGGAATATTTACTGTGGAA	2697
Db	2581	TTTGAAGGCATCTACACCATTAAGAGTGTCTGCTCATATGGAATATTTACTGTGGAA	2640
Qy	2698	ATCTTCTCACTTGTGTGAATCCTTACCCCTGGCATTCGGTTGATGCTAACTTCTACAAA	2757
Db	2641	ATCTTCTCACTTGTGTGAATCCTTACCCCTGGCATTCGGTTGATGCTAACTTCTACAAA	2700
Qy	2758	CTGAATCAAAATGGAATTTAAATGATCAGCCATTTTATGCTTACAGAGAAATATACATT	2817
Db	2701	CTGAATCAAAATGGAATTTAAATGATCAGCCATTTTATGCTTACAGAGAAATATACATT	2760
Qy	2818	ATAATGCAATCCTGCTGGGCTTTGACTCAAGGAAACGGCCATCCTTCCCTAATTTGACT	2877
Db	2761	ATAATGCAATCCTGCTGGGCTTTGACTCAAGGAAACGGCCATCCTTCCCTAATTTGACT	2820
Qy	2878	TCGTTTTTAGGATGTACGCTGGCAGATGCAGAGAAAGCGATGTATCAGAAATGTGGATGSC	2937
Db	2821	TCGTTTTTAGGATGTACGCTGGCAGATGCAGAGAAAGCGATGTATCAGAAATGTGGATGSC	2880
Qy	2938	CGTGTTCGGAATGTCTCTCACACCTTACCAAAACAGGGCCCTTTTCAGCAGAGATGGAT	2997
Db	2881	CGTGTTCGGAATGTCTCTCACACCTTACCAAAACAGGGCCCTTTTCAGCAGAGATGGAT	2940
Qy	2998	TTGGGGCTACTCTCTCCGAGGCTCAGTCAAGATTCGTAGAGAAACAATTTAGTTTTTA	3057
Db	2941	TTGGGGCTACTCTCTCCGAGGCTCAGTCAAGATTCGTAGAGAAACAATTTAGTTTTTA	3000
Qy	3058	AGGACTTCATCCCTCCACCTATCCCTAACAGGCTGTAGATTACCAAAACAGATTAATTT	3117
Db	3001	AGGACTTCATCCCTCCACCTATCCCTAACAGGCTGTAGATTACCAAAACAGATTAATTT	3060
Qy	3118	CATCACTAAAGAAATCTATTTATCACTGCTGCTTCCACAGACTTTTCTCTAGAAGCCG	3177
Db	3061	CATCACTAAAGAAATCTATTTATCACTGCTGCTTCCACAGACTTTTCTCTAGAAGCCG	3120

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2003, 07:20:58 ; Search time 33.0166 Seconds
(without alignments)
2892.346 Million cell updates/sec

Title: US-09-919-408A-4

Perfect score: 5274

Sequence: 1 MPALARDAGTVPLLVVFSAM.....PFSREMDLGLSPQAQVEDS 993

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 76:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	5238	99.3	993	2 A36873	protein-tyrosine k
2	4533.5	86.0	1000	2 S18827	Flt3 protein - mou
3	4429.5	84.0	992	2 A39931	protein-tyrosine k
4	1286	24.4	977	2 I45877	protein-tyrosine k
5	1266	24.0	975	1 TVMSKT	protein-tyrosine k
6	1263.5	24.0	978	1 A49814	protein-tyrosine k
7	1251	23.7	976	1 TVHUKT	protein-tyrosine k
8	1226	23.2	954	2 I51703	c-kit-related kina
9	1225.5	23.2	980	1 TVCTMD	macrophage colony-
10	1218	23.1	941	1 TVVMVD	protein-tyrosine k
11	1216.5	23.1	972	1 TVHUMD	macrophage colony-
12	1215	23.0	960	1 JN0677	protein-tyrosine k
13	1213	23.0	975	2 T30816	macrophage colony-
14	1194	22.6	978	2 S16385	macrophage colony-
15	1183.5	22.4	976	1 TVMSMD	macrophage colony-
16	1176	22.3	1088	1 PFRTGA	macrophage colony-
17	1166	22.1	1089	1 PFHUGA	platelet-derived g
18	1146	21.7	1089	1 S33727	platelet-derived g
19	1128	21.4	1087	2 I51552	platelet-derived g
20	1105.5	21.0	1098	1 PFMSRB	platelet-derived g
21	1086.5	20.6	1106	1 PFHUGB	platelet-derived g
22	1080	20.5	1048	2 T30815	platelet-derived g
23	992	18.8	790	1 FOMVHZ	Sag-kit polypeptid
24	990	18.8	1338	2 S09882	protein-tyrosine k
25	959.5	18.2	1333	2 I78875	receptor tyrosine
26	956	18.1	1336	2 I60598	Fit-1 tyrosine kin
27	944.5	17.9	1356	2 JCI402	protein-tyrosine k
28	939.5	17.8	1330	2 S49010	embryonic receptor
29	935	17.7	1348	2 S51656	vascular endotheli

30	932	17.7	1379	2 JC4954	vascular endotheli
31	923	17.5	1367	2 A41228	protein-tyrosine k
32	907.5	17.2	1363	2 I58375	protein-tyrosine k
33	904.5	17.2	1298	2 A48999	protein-tyrosine k
34	874.5	16.6	823	2 B35963	protein-tyrosine k
35	851	16.1	160	2 A39061	protein-tyrosine k
36	849	16.1	821	1 TVHUF2	fibroblast growth
37	842	16.0	813	1 A49123	fibroblast growth
38	841	15.9	824	2 S24108	protein-tyrosine k
39	836.5	15.9	822	2 A45081	fibroblast growth
40	832.5	15.8	824	2 S36439	fibroblast growth
41	828.5	15.7	769	2 S16236	fibroblast growth
42	828.5	15.7	822	2 A41794	keratinocyte growt
43	827	15.7	797	2 S38579	fibroblast growth
44	825	15.6	821	1 TVMSBK	fibroblast growth
45	823.5	15.6	822	2 B54846	fibroblast growth

ALIGNMENTS

RESULT 1

A36873 protein-tyrosine kinase (EC 2.7.1.112) STK-1 precursor - human
N:Alternate names: stem cell tyrosine kinase 1
C:Species: Homo sapiens (man)
C>Date: 03-Nov-1995 #sequence_revision 03-Nov-1995 #text_change 04-Feb-2000
C/Accession: A36873
R/Small, D.; Levenstein, M.; Kim, E.; Carow, C.; Amin, S.; Rockwell, P.; Witte, L.; Bur
Proc. Natl. Acad. Sci. U.S.A. 91, 459-463, 1994
A>Title: STK-1, the human homolog of Plk-2/Flt-3, is selectively expressed in CD34(+) h
A/Reference number: A36873; MUID:94119906; PMID:7507245
A/Accession: A36873
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-993 <SMA>
A/Cross-references: GB:U02687
A/Note: in the authors translation, an additional residue Ala is shown after 420-Ala an
C/Genetics:
A/Map position: 13q12
C/Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homol.
C/Keywords: ATP; glycoprotein; phosphotransferase; transmembrane protein; tyrosine-spec
F/608-950/Domain: protein kinase homology <Kin>
F/616-624/Region: protein kinase ATP-binding motif

Query Match 99.3%; Score 5238; DB 2: Length 993;
Best Local Similarity 99.7%; Pred. No. 4.2e-248;
Matches 991; Conservative 0; Mismatches 1; Indels 2; Gaps 2;

QY	1	MPALARDAGTVPLLVVFSAMIFGTITNQDLFVVKVILNHKNNDSSVYKSSSSYPMVSESP	60
DB	1	MPALARDAGTVPLLVVFSAMIFGTITNQDLFVVKVILNHKNNDSSVYKSSSSYPMVSESP	60
QY	61	EDLGCALRQSSGTVEAAAEVDVVSASITQLVLDVAPGNISCLVWPKHSLNCQPHFDL	120
DB	61	EDLGCALRQSSGTVEAAAEVDVVSASITQLVLDVAPGNISCLVWPKHSLNCQPHFDL	120
QY	121	QNRGVSWILKMTETQAGEYLLFIQSEATNYTLFTVSIIRNTLLYTLRPPYFRKMENQD	180
DB	121	QNRGVSWILKMTETQAGEYLLFIQSEATNYTLFTVSIIRNTLLYTLRPPYFRKMENQD	180
QY	181	ALVCISESPEPIVEWVLCDSQGESCKESPAVVKKEKVLHFLFGTDIRCCARNELGRE	240
DB	181	ALVCISESPEPIVEWVLCDSQGESCKESPAVVKKEKVLHFLFGTDIRCCARNELGRE	240
QY	241	CTRLFTIDLNQTPQTLPOLFLKVGEPFLMIRCAVHVNHGFGLTWLENKALBEGNYFEM	300
DB	241	CTRLFTIDLNQTPQTLPOLFLKVGEPFLMIRCAVHVNHGFGLTWLENKALBEGNYFEM	300
QY	301	STYSTNRTMIRILFAFVSSVARNDTGYTCSSSKHPSQSALVTIVGKGFNATNSSDYE	360
DB	301	STYSTNRTMIRILFAFVSSVARNDTGYTCSSSKHPSQSALVTIVGKGFNATNSSDYE	360

QY 361 IDQYEFCEPSVRKAYPOIRCTWTFRRKSPCEQKGLDNGYSISKPNHKGHPGEYIEH-- 419
 DB 361 IDQYEFCEPSVRKAYPOIRCTWTFRRKSPCEQKGLDNGYSISKPNHKGHPGEYIEH 420
 QY 420 AENDDAQFTKMFNTNRRKPOVLAEASASQSCFSDGYPLPSMTWKKCDKSPNCTEEIT 479
 DB 421 AENDDAQFTKMFNTNRRKPOVLAEASASQSCFSDGYPL-SMTWKKCDKSPNCTEEIT 479
 QY 480 EGVNWKANRKFQGVQWSSSTLNMSBAIKGFLVKCCAYNSLGTSCETILLNSGPPFPFIQ 539
 DB 480 EGVNWKANRKFQGVQWSSSTLNMSBAIKGFLVKCCAYNSLGTSCETILLNSGPPFPFIQ 539
 QY 540 DNISFYATTGCVLLFIWLTLLIHKYKQFRYSQLOQVQVGTSSDNEFYVDREY 599
 DB 540 DNISFYATTGCVLLFIWLTLLIHKYKQFRYSQLOQVQVGTSSDNEFYVDREY 599
 QY 600 DLKWEFPRENLEFGKVLGSAFGKVMNATAYGSKTGVSIOQAVKMLKEKADSSERELM 659
 DB 600 DLKWEFPRENLEFGKVLGSAFGKVMNATAYGSKTGVSIOQAVKMLKEKADSSERELM 659
 QY 660 SELKMMTQLGSHENIENVLLGACTLSGPYLIIFYCCYGDLLNLYLSKREKPHRTWTEIFK 719
 DB 660 SELKMMTQLGSHENIENVLLGACTLSGPYLIIFYCCYGDLLNLYLSKREKPHRTWTEIFK 719
 QY 720 EHNFSFYPTQSHPNSSMPGSRVQIHPDSDQISGLHNSFHSDEIEYENQKLEBEED 779
 DB 720 EHNFSFYPTQSHPNSSMPGSRVQIHPDSDQISGLHNSFHSDEIEYENQKLEBEED 779
 QY 780 LNVLTTFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVKICDPGLARDIMSD 839
 DB 780 LNVLTTFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVKICDPGLARDIMSD 839
 QY 840 SNYVVRGNARLPVKWMAPESLFEGITIKSDVMSYGILLWEIFSLGVNYPGIPVDANFY 899
 DB 840 SNYVVRGNARLPVKWMAPESLFEGITIKSDVMSYGILLWEIFSLGVNYPGIPVDANFY 899
 QY 900 KLTONGFMDQPPYATEEIIYIMQSWAFDSRKRPSFPNLTSLFLGCQLADAEAMYNQVD 959
 DB 900 KLTONGFMDQPPYATEEIIYIMQSWAFDSRKRPSFPNLTSLFLGCQLADAEAMYNQVD 959
 QY 960 GRVSECPHTYQNRPPSREMDLGLLSPQAQVEDS 993
 DB 960 GRVSECPHTYQNRPPSREMDLGLLSPQAQVEDS 993

RESULT 2
 S18827
 Flt3 protein - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 19-Feb-1994 #sequence_revision 10-Nov-1995 #text_change 24-Sep-1999
 C:Accession: S18827
 R:Rosnet, O.; Marchetto, S.; deLapeyriere, O.; Birnbaum, D.
 Oncogene 6, 1641-1650, 1991
 A:Title: Murine Flt3, a gene encoding a novel tyrosine kinase receptor of the PDGFR/CSF1
 A:Reference number: S18827; MUID:92019834; PMID:1656368
 A:Accession: S18827
 A:Molecule type: mRNA
 A:Residues: 1-1000 <ROS>
 A:Cross-references: EMBL:X59398; NID:G50978; PIDN:CAA42041.1; PID:G50979
 C:Superfamily: unassigned Ser/thr or Tyr-specific protein kinases; protein kinase homolo
 F:609-953/Domain: protein kinase homology <KIN>
 F:617-625/Region: protein kinase ATP-binding motif

Query Match 86.08; Score 4533.5; DB 2; Length 1000;
 Best Local Similarity 85.64; Pred. No. 9.3e-214;
 Matches 851; Conservative 56; Mismatches 84; Indels 3; Gaps 2;
 QY 1 MPALA-RDAGTVLLVVFSAIFGTITNQDLPVTKCVLINHKNDSSGVKSSSYPMPVSES 59
 DB 1 MRALAQRSDRLLLLVLSWILETVNQDLPVTKCVLISHENNGSAGKPSRWVRGS 60
 QY 60 PEDLGCALRPQSSGTVYAAAVEVDVASITLQVLDPAGNISCLWVFKHSLNCQPHFD 119

DB 61 PEDLQCTPRQSEGTVEAATVEAESGSITLQVQLATPGDLSCLWFKHSLLCQPHFD 120
 QY 120 LQNRGVVSVLLKMTTQAGEYLLFTQSEATNTTILFTVISIRWTLTYTLRRPFRKXENQ 179
 DB 121 LQNRGVISMAILNVTETQAGEYLLHTQSEAAANTVLTFTVNRDQTQVLYLRPFRKXENQ 180
 QY 180 DALVCISESVPEPIVEMVLCDQSECKEESPAVVKKEKVLHFGTDRCCARNELGR 239
 DB 181 DALLCISEGVPEPTEVWVLCSSHRECKEESPAVVRKEKVLHFGTDRCCARNALGR 240
 QY 240 ECTRLFTIDLNTPOPTTLPOLFLKVGEPWIRKCAVHVNHGFGLTWELENKALEEGNYFE 299
 DB 241 ESTKLTIDLNQAPQSTLPOLFLKVGEPWIRKCAIHVNHGFGLTWELEDKALEEGSYFE 300
 QY 300 MSTYSTNRTMIRLFAFVSSVARNDGYTTCSSSKHPSQSALVTIIVKGFINATNSSEDI 359
 DB 301 MSTYSTNRTMIRILLAFVSSVGRNDGYTTCSSSKHPSQSALVTILEKGINATSSQEEY 360
 QY 360 EIDQYEFCEPSVRKAYPOIRCTWTFRRKSPCEQKGLDNGYSISKPNHKGHPGEYIEH 419
 DB 361 EIDPYEKFCEPSVRKAYPRIRCTWTFRRKSPCEQKGLDNGYSISKPNHKGHPGEYIFY 420
 QY 420 AENDDAQFTKMFNTNRRKPOVLAEASASQSCFSDGYPLPSMTWKKCDKSPNCTEEIT 479
 DB 421 AENDDAQFTKMFNTNRRKPOVLAEASASQSCFSDGYPLPSMTWKKCDKSPNCTEEIP 480
 QY 480 EGVNWKANRKFQGVQWSSSTLNMSBAIKGFLVKCCAYNSLGTSCETILLNSGPPFPFIQ 539
 DB 481 EGVNWKANRKFQGVQWSSSTLNMSBAIKGFLVKCCAYNSMGTSCTIFLNSGPPFPFIQ 540
 QY 540 DNISFYATTGCVLLFIWLTLLIHKYKQFRYSQLOQVQVGTSSDNEFYVDREY 599
 DB 541 DNISFYATTGCVLLFIWLTLLIHKYKQFRYSQLOQVQVGTSSDNEFYVDREY 600
 QY 600 DLKWEFPRENLEFGKVLGSAFGKVMNATAYGSKTGVSIOQAVKMLKEKADSSERELM 659
 DB 601 DLKWEFPRENLEFGKVLGSAFGKVMNATAYGSKTGVSIOQAVKMLKEKADSSERELM 660
 QY 660 SELKMMTQLGSHENIENVLLGACTLSGPYLIIFYCCYGDLLNLYLSKREKPHRTWTEIFK 719
 DB 661 SELKMMTQLGSHENIENVLLGACTLSGPYLIIFYCCYGDLLNLYLSKREKPHRTWTEIFK 720
 QY 720 EHNFSFYPTQSHPNSSMPGSRVQIHPDSDQISGLHNSFHSDEIEYENQKLEBEED 779
 DB 721 EHNFSFYPTQSHPNSSMPGSRVQIHPDSDQISGLHNSFHSDEIEYENQKLEBEED 780
 QY 778 EDLNVLTTFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVKICDPGLARDIM 837
 DB 781 EDLNVLTTFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVKICDPGLARDIL 840
 QY 838 SDSNYVVRGNARLPVKWMAPESLFEGITIKSDVMSYGILLWEIFSLGVNYPGIPVDAN 897
 DB 841 SDSNYVVRGNARLPVKWMAPESLFEGITIKSDVMSYGILLWEIFSLGVNYPGIPVDAN 900
 QY 898 FYKLQNGFMDQPPYATEEIIYIMQSWAFDSRKRPSFPNLTSLFLGCQLADAEAMYNQ 957
 DB 901 FYKLQNGFMDQPPYATEEIIYIMQSWAFDSRKRPSFPNLTSLFLGCQLADAEAMYNQ 960
 QY 958 VGRVSECPHTYQNRPPSREMDLGLLSPQAQVE 991
 DB 961 MCGNVPEHPSIYQNRPPSREMDLGLLSPQAQVK 994

RESULT 3
 A39931
 protein-tyrosine kinase (EC 2.7.1.112) flk-2, hematopoietic - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 17-Jan-1992 #sequence_revision 17-Jan-1992 #text_change 04-Feb-2000
 C:Accession: A39931
 R:Matthews, W.; Jordan, C.T.; Wiegand, G.W.; Pardoll, D.; Lemischka, I.R.
 Cell 65, 1143-1152, 1991
 A:Title: A receptor tyrosine kinase specific to hematopoietic stem and progenitor cell-

A:Reference number: A39931; MUID:91292518; PMID:1648448

A:Accession: A39931

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-992 <NAT>

C:Cross-references: GB:M64689; NID:9193327; PIDN:AAA37634.1; PID:9193328

C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog

C:Keywords: ATP; phosphotransferase; transmembrane protein; tyrosine-specific protein ki

F:609-953/Domain: protein kinase homolog <KIN>

F:617-625/Region: protein kinase ATP-binding motif

Query Match 84.0%; Score 4429.5; DB 2; Length 992;

Best Local Similarity 84.1%; Pred. No. 1.1e-208;

Matches 836; Conservative 57; Mismatches 90; Indels 11; Gaps 4;

QY 1 MPALA-RDAGTVPLLVFAMIFGTTNQDLPVVKVILNHKNDDSSVSKSSYPWVSS 59
 DB 1 MEALQSRDRLLLVLSVMIETVTNQDLPVVKVILSHENNGSSAGKPSYMRGS 60
 QY 60 PEDLGCALRPOSGTVYEAADVVSASITLQVLVDAPGNISCLWVFKHSSINCOHPED 119
 DB 61 PEDLOCTPRQSGTVYEAADVVSASITLQVLVDAPGNISCLWVFKHSSINCOHPED 120
 QY 120 LQNRGVSVMLKMTQAGEYLLFTQSEATNTYILFTVSIIRNTLYTLRRPYFRKMNQ 179
 DB 121 LQNRGIVSNAILNVETQAGEYLLHLOSERANYTLFTVNRDQTLVLRPYFRKMNQ 180
 QY 180 DALVCISEVPPIVENVLDCSGESCKESPAVKKKEKVLHELPGTDIRCARNELGR 239
 DB 181 DALLCISEGPEPTVENVLDCSHRESCKESGPAVKKKEKVLHELPGTDIRCARNALGR 240
 QY 240 ECTRFTIDLNQPTTLPOLFLKVGEPVIRCKAVHNVHGFGLTWELNKALEEGNYE 299
 DB 241 ECKTFTIDLNAOPOSTLPOLFLKVGEPVIRCKAVHNVHGFGLTWELNKALEEGNYE 300
 QY 300 MSETYSTNRTMIRILPAFVSVARNDGYTTCSSSKHPSQSALVTIVKGFINATNSSEY 359
 DB 301 MSETYSTNRTMIRILPAFVSVARNDGYTTCSSSKHPSQSALVTILEKGFINATNSSEY 360
 QY 360 EIDQVEEFCFVRKAYPOIRCTWTFSRKSPFCEQKGLDNGYSISKFCNHKHPGEYIFH 419
 DB 361 EIDPEKFCFVRKAYPRICTWTFSPQASFPCEQKGLDNGYSISKFCNHKHPGEYIFH 420
 QY 420 AENDDAQFTMTLNIIRKPOVLAESAQSCFSDGYPLPSWTWKCDKSPNCTEET 479
 DB 421 AENDDAQFTMTLNIIRKPOVLAESAQSCFSDGYPLPSWTWKCDKSPNCTEET 480
 QY 480 EGVWNRKANRVFGQWVSSSTLNMSAIGKFLVCCAYNSLGTSCETILLNSPGPPFIQ 539
 DB 481 EGVWNRKANRVFGQWVSSSTLNMSAIGKFLVCCAYNSLGTSCETILLNSPGPPFIQ 540
 QY 540 DNISFYATIGVCLLFIVVLTLIHKYKQFRYESQLOMVQVGTSSDNEYFYVDREY 599
 DB 541 DNISFYATIGVCLLFIVVLTLIHKYKQFRYESQLOMVQVGTSSDNEYFYVDREY 600
 QY 600 DLKWFPPRENLFKGLVGSAGFQVKNATAYGSKTGVSIQVAKMLKBAKDSSEALM 659
 DB 601 DLKWFPPRENLFKGLVGSAGFQVKNATAYGSKTGVSIQVAKMLKBAKDSSEALM 660
 QY 660 SELKMTQLGSHENIVNLGACTLSPIVLIPEYCCYGLLNVLRSKRKHRTWTEIFK 719
 DB 661 SELKMTQLGSHENIVNLGACTLSPIVLIPEYCCYGLLNVLRSKRKHRTWTEIFK 720
 QY 720 ENHFSYPTFQGHNSMPSREVQTHPDSDOIQLHNSFSEDEIEYENOKRL--EE 777
 DB 721 ENHFSYPTFQGHNSMPSREVQTHPDSDOIQLHNSFSEDEIEYENOKRLAEE 780
 QY 778 EDNLVLTFFEDLLCFAYQVAKMEFLFKSCVHRDLAARNVLTGHKVKICDFGLARDIM 837
 DB 781 EDNLVLTFFEDLLCFAYQVAKMEFLFKSCVHRDLAARNVLTGHKVKICDFGLARDIL 840
 QY 838 SDSNVVVRGNARLPVKWMAPESLFEGIYTIKSDVMSYGILLWEIFSLGVNPNPFIQ 897

DB 841 SDSNVVVRGNARLPVKWMAPESLFEGIYTIKSDVMSYGILLWEIFSLGVNPNPFIQ 900
 QY 898 FYKLQNGFKMDOPFYATEEIIYIMOSWAPSRKRPSPNLTSLGCOLADAEEMYQN 957
 DB 901 FYKLQNGFKMDOPFYATEEIIYIMOSWAPSRKRPSPNLTSLGCOLADAEEMYQN 957
 QY 958 VDGVRSECPHTYQNRPRPSREMDLGLLSPQAQVE 991
 DB 958 ----TFTSHLPKQAPQORG-GLRAQSPORQVK 986
 RESULT 4
 I45877
 protein-tyrosine kinase (EC 2.7.1.112), receptor type kit precursor - aurochs
 C:Species: Bos primigenius (aurochs)
 C:Update: 19-Dec-1997 #sequence revision 19-Dec-1997 #text_change 20-Apr-2000
 C:Accession: I45877
 R:Kubota, T.; Hikono, H.; Sasaki, E.; Sakurai, M.
 Gene 141, 305-306, 1994
 A:Title: Sequence of a bovine c-kit proto-oncogene cDNA.
 A:Reference number: I45877; MUID:94215924; PMID:7512939
 A:Accession: I45877
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-977 <KUB>
 A:Cross-references: GB:D16680; NID:9516659; PIDN:BA04084.1; PID:9516660
 C:Superfamily: macrophage colony-stimulating factor 1 receptor; immunoglobulin homology.
 C:Keywords: ATP; phosphotransferase; tyrosine-specific protein kinase
 F:329-395/Domain: immunoglobulin homology <IMM>
 F:588-932/Domain: protein kinase homology <KIN>
 Query Match 24.4%; Score 1286; DB 2; Length 977;
 Best Local Similarity 33.2%; Pred. No. 1.8e-55;
 Matches 332; Conservative 172; Mismatches 344; Indels 152; Gaps 36;
 QY 47 VCKSSYPWVSSPDLGC-ALRPOSGTVYEAADVVSASITLQVLVDAPGNISCLW 105
 DB 20 VOTGSSQPSV--SPGELSLPSHPAKS-----ELIVSGDEIRLCTDPGFVK--W 66
 QY 106 VFKHSLNLCQPHFDLQNRGWSMVILKMTQAGEYLLFIQSEAT---NYTLFTVSI 162
 DB 67 TFE-----ILGQLSEKTNPEWIT-EKAEATNTGNTYCTNKGGLSS 105
 QY 163 TL-----LYTLRRPYFRKMNQDALVCISEVPPIVENVLDCSGESCKESPAV 213
 DB 106 SIYFVRDPEKFLDIDPLYGKEEN-DTLVRCPLDPE-VTNYSLTGCEGKPLKDLTFV 163
 QY 214 -----VKKEKVLHELPGTDIRC--CARNELGRE-CTRLFTID----LNQTPQTLLP 258
 DB 164 ADPKAGITRNVRKYHRL-----CLHCSANQRKSLSKKFTLKVRAAIKAVPVVSVS 217
 QY 259 Q--LPLKVGEPVIRCKAVHNVHGFGLTWELNKALEEGNYPEMSTYSTNRTMIRILFAF 316
 DB 218 KTSYLLRGEERFAVTLIKDVSSVDSMWIKENSQOTKATQKNSWHQDGFYLRERLT 277
 QY 317 VSSVARNDGYTTCSSSKHPSQSALVT---IVKGFINA-----TNSSEDEYIDQY 364
 DB 278 ISSARVNDSGVPMCVANNTFGSANVTTLVVVDKGFINIFPMWNTTVFVNDGENVDL--- 334
 QY 365 EEPFCFVRKAYPO-IRCTWTFSRKS-----PCEQKGLDNGYSISKFCNHK 411
 DB 335 -----VVEYEAPKVRQWYIMNRTSTDKWDYDPKSE-----NESNIRYVNLHLTRLK 384
 QY 412 --OPGEYIFHAENDDAQFTMTLNIIRKPOVLAESA--SQASCFSDGYPLPSMTWK 467
 DB 385 GTEGTYTFHVSNSDVNSVTFNVVNTKPELITDRLVNGMLQCVAAAGFPETIDWYFC 444
 QY 468 SDKSPNCTEETI--TEGVWNRKANRVFGQWVSSSTLNMSAIGKFLVCCAYNSLGTSCET 536
 DB 445 POTEQRCSVPVGVDPVQIQNSSVSPFKLVVYSTIDDDSTFKHNGTVECRAYNDVGKSSAS 504
 QY 527 ILLNSPGPPF-----FIQDNISFYATIGVCLLFIVVLTLIHKYKQFRYESQLOM 578

505 FNFAPKSGNSKEQIHAHTLETPILLIGFVIAAGLMCIPIVMILT-----YKYLQKPMYEVQWKV 560
 579 V-QVTGSSDNEYFYDFREYEDKWEFFPRENLEFGKVLGSGAFKGMWNTATYIGSKTV 637
 561 VEEING-----NNVYIDPTQLPYDHRKWEFFPRNLSFGKTLGAGAFKGVVEATYGLIKSDA 617
 638 SIQVAVKMLKEVADSSERALSSELKMTQOLGSHENIVNLLGACTLSGPIYILFEVCCYG 697
 618 AMTAVKMLKPSAHLTEREALMSKLVSLYGLNHNMINVLLGACTTGGTTLVITEYCCYG 677
 698 DLLNVLRSKREKFRHTWTS-----IFKEHNSFYPTQSHPNSSM---PGSREVQIHPD 748
 678 DLNLEFRKRSDFISCKQEDHAEVALYKMLLSKSSCNDSTNEYMDKPGVSYV-VPTK 736
 749 SQOISGLHNSFHSDEIEYENQKLEEDNLVTFEDLLCFAYQVAKGMFLEPKSCV 808
 737 ADKRSRARSY-----IERDVTPIAIMEDEL-ALDLELLSFSYQVAKGMFLASKNCI 790
 809 HRDLAARNVLTHGKVKKICDFGLARDIMSDSNVYVGRNARLPVKWMAPESEPIYTIK 868
 791 HRDLAARNILLTHGRITKICDFGLARDIKNSNVYVVGKRNARLPVKWMAPESEPCVYTFE 850
 869 SDVWSYGILLWEIPEFSGVNPYPCIPVDANFYKLIONGKMDQPFYATEEIIYIMQSCWAF 928
 851 SDVWSYGIFLWELFSLGSSPYGMPVDSKPYKMKEGFRMLSPHAPAEYIMDKTCWDA 910
 929 DSKRSPFNLSFLGCLQADAEAMYNQVNDGRVSEC-PH 967
 911 DPLKRTFTKQIVQLIEKQISESTNHYSN-----LANCSPH 946
 RESULT 5
 TVMSKT
 N:Alternate names: tyrosine kinase receptor c-kit
 C:Species: Mus musculus (house mouse)
 C:Date: 31-Mar-1991 #sequence revision 31-Mar-1991 #text_change 31-Mar-2000
 C:Accession: S00474; B44876; I49596
 R:Qiu, F.; Ray, P.; Brown, K.; Barker, P.E.; Jhanwar, S.; Ruddle, F.H.; Besmer, P.
 EMBL J. 7, 1003-1011, 1988
 A:Title: Primary structure of c-kit: relationship with the CSF-1/PDGF receptor kinase fa
 A:Reference number: S00474; MUID:88296403; PMID:2456920
 A:Accession: S00474
 A:Molecule type: mRNA
 A:Residues: 1-975 <QIU>
 A:Cross-references: GB:Y00864; NID:G50423; PIDN:CAA68772.1; PID:G50424
 R:Rossi, P.; Marziani, G.; Albanesi, C.; Charlesworth, A.; Geremia, R.; Sorrentino, V.
 Dev. Biol. 152, 203-207, 1992
 A:Title: A novel c-kit transcript, potentially encoding a truncated receptor, originates
 A:Reference number: A44876; MUID:92331813; PMID:1378413
 A:Accession: B44876
 A:Molecule type: DNA
 A:Residues: 771-814 <ROS>
 A>Note: sequence extracted from NCBI backbone (NCBI:108837, NCBI:108840)
 R:Yasuda, H.; Galli, S.J.; Geisler, E.N.
 Biochem. Biophys. Res. Commun. 191, 893-901, 1993
 A:Title: Cloning and functional analysis of the mouse c-kit promoter.
 A:Reference number: I49596; MUID:93221533; PMID:7682073
 A:Accession: I49596
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-22 <RES>
 A:Cross-references: GB:L11358; NID:G293325; PIDN:AAA37420.1; PID:G293326
 C:Genetics:
 A:Gene: kit; c-kit
 A:Map position: 5
 C:Superfamily: macrophage colony-stimulating factor 1 receptor; immunoglobulin homology;
 C:Keywords: APP; autophosphorylation; glycoprotein; kinase-related transforming protein;
 fic protein kinase
 F:1-22/Domain: signal sequence #status predicted <SIG>
 F:23-975/Product: protein-tyrosine kinase kit #status predicted <KTC>
 F:23-519/Domain: extracellular #status predicted <EXT>
 F:51-100/Domain: immunoglobulin homology <IMM1>

F:130-189/Domain: immunoglobulin homology <IMM2>
 F:221-295/Domain: immunoglobulin homology <IMM3>
 F:331-397/Domain: immunoglobulin homology <IMM4>
 F:426-496/Domain: immunoglobulin homology <IMM5>
 F:520-542/Domain: transmembrane #status predicted <TMM>
 F:543-975/Domain: intracellular #status predicted <INT>
 F:586-929/Domain: protein kinase #status predicted <KIN>
 F:594-602/Region: protein kinase ATP-binding motif
 F:58-98,137-187,234-293,431-494/Disulfide bonds: #status predicted
 F:146,296,303,323,355,370,466,489/Binding site: carboxyrate (Asn)
 F:622,639,790/Active site: Lys, Glu, Asp #status predicted
 F:795,808/Binding site: magnesium (Asn, Asp) #status predicted
 Query Match 24.0%; Score 1266; DB 1; Length 975;
 Best Local Similarity 33.2%; Pred. No. 1.7e-54;
 Matches 336; Conservative 173; Mismatches 330; Indels 172; Gaps 41;
 QY 35 CVLIHNKNDSSVGKSSYPMVSEPEDLGALRQSSGTVYEAATAVEVDVSASITLQVL 94
 DB 12 CVLLVLLRGQTATSPASPGEPSP-----SIHPAQS-----ELIVEAGDTL--- 54
 QY 95 VDAPGNISCL-----WVPKHSLLNCQPHFDLQNRGWSMVLKMTETQAGYLLFIQSE 148
 DB 55 -----SLTICIDPDFVRWTFK-----TYFN-----EMVENKKNEWIQ-EKAE 89
 QY 149 ANYTILFTVSRNLT-----LYTLRPYFRKMNODALVCISSVPEPIVE 195
 DB 90 ATR-TGTYTCSNSNGLTSSIVYFVRDPAKFLVGLPLFGK-EDSDALVRCPLTDPQ-VSN 146
 QY 196 WYLCDSQGESCKEESPAV-----VKKEEVKLHELFGTDIRCCARNELGRECTRL--- 244
 DB 147 YSLIEDCGKSLPTDLTFVNPKNAGITIKNVKAYKRLC---VRCAAQ-----RDGTWLHSD 199
 QY 245 -FTIDINOT-----POTLLPOL--FLKVGEPFLMIRKAVHNHGFGLTW-----E 286
 DB 200 KTLKVRRAIKAIPIVVPVETSHLLKKGDTFTVCTIKDVSTSVNMLKMNPPQOHIAQ 259
 QY 287 LENKALBEGNYEMSTYTNRMIRILFAFVSGVARNDTGYTSCSSKHPQSALVT--- 343
 DB 260 VKHNSWHRGDF---NYERQETLT-----ISSARVDDSGVFMVCYANNTFGSANVTTLK 309
 QY 344 IVYKGFINATN--SSEYEDIDVEEFCFSVRFAKYPQ-IRCTWTFSRKSPCEQKGLD--- 398
 DB 310 VVEKGFINISPVKNTFTVFDGENVDLVVEYKAYKPEHQOQIYMRT---SANKGDIYK 367
 QY 399 -NGYSISKFCNKH-----OPGEYIHAENDDAQFTKMTLINIRKPOVLA--EASAS 448
 DB 368 SDNKSNIYVNLRLTLKGTGGTYTFLVNSDSASVTFNVVNTKPELLTYDRLING 427
 QY 449 QASCFSDDGYPLPSWTWKCKSDKSPNCTEITE-GWNRKANRKVFGQWSSSTLNMSEAI 507
 DB 428 MLQCVAEAGPEPTIDWYFCTGAEQRCTTPVSPVDVQVNVSVSPFGKLVQSSIDSVPFR 487
 QY 508 KGFLVKCCAYNSLGTSCETILLNSPGPEFP---IQDN-----ISFYATIGVCLLFIV 557
 DB 488 HNGTVECKASNDVGKS--SAFFN-----FAFKQIQIAHTLFTPLLLIGFVVAAGANGIIVM 541
 QY 558 LTLILCHYKQKQFRYESQQLQW-QVTGSSDNEYFYVDREYEDLKWEPFRENLEFGKVL 616
 DB 542 LT-----YKYLQPMYEVQWKVVEING---NNVYIDPTQLPYDHRKWEFPRNLSFGKTL 594
 QY 617 GSGAFGKVMNATYIGSKTVGIQVAVKMLKEKADSSEREALMSKMTQOLGSHENIVN 676
 DB 655 LLGACTLGGTTLVITEYCCYGLNLRKRRKRSFIFSKQEQEAALYKMLHSTFESCD 714
 QY 731 SHPNSSM---PGSREVQIHPDSDQISGLHNSFHSDEIEYENQKLEEDNLVTFED 787
 DB 715 S-SNEYMDKPGVSYV-VPTKTDKRSARIDSY-----IERDVTPIAIMEDEL-ALDLD 766

788 LLCFAVOAKGMEFLFKSCVHRDLAARNVLVTHGKVKICDFGLARDIMSDSNVYVRGN 847
 767 LLSFSQVAKAFKASKNCIHRDLAARNLLTHGRTKICDFGLARDIMSDSNVYVRGN 826
 848 ARLPVKMAPELFGIYTIKSDVWSYIGLLWEIFSLGVNYPGIPVDANFYKLIQNGFK 907
 827 ARLPVKMAPELFGIYTIKSDVWSYIGLLWEIFSLGVNYPGIPVDANFYKLIQNGFK 886
 908 MDQPFVATEIYIIMOSCHAFDSRKPSPNLTSLFGCLADAEAEAMQNV 958
 887 MVSPEHAPAEYDMVTKDADPLKPTKQVQVQLIEKQISDSTKHYSNL 937

RESULT 6
 A49814
 protein-tyrosine kinase (EC 2.7.1.112), receptor type kit precursor - rat
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 16-Jun-2000
 C:Accession: A49814; S49088
 R:Tsujiyama, T.; Hirota, S.; Nomura, S.; Niwa, Y.; Yamazaki, M.; Tono, T.; Morii, E.; Kikuchi, T.; Tsujiyama, T.; Tono, T.; Yamazaki, M.; Nomura, S.; Kitamura, Y.
 Blood 78, 1942-1946, 1991
 A:Title: Characterization of Ws mutant allele of rats: a 12-base deletion in tyrosine kinase
 A:Reference number: A49814; MUID:92003944; PMID:1912577
 A:Accession: A49814
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-978 <TSU>
 A:Cross-references: GB:D12524; NID:g220707; PIDN:BA02094.1; PID:g220708
 R:Tsujiyama, T.; Tono, T.; Yamazaki, M.; Nomura, S.; Kitamura, Y.
 submitted to the EMBL Data Library, October 1991
 A:Description: Two isoforms of rat c-kit receptor tyrosine kinase.
 A:Reference number: S49088
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-511, 516-978 <TSU>
 A:Cross-references: EMBL:X62491; NID:g509135; PIDN:CAA44354.1; PID:g509136
 C:Superfamily: macrophage colony-stimulating factor 1 receptor; immunoglobulin homology;
 C:Keywords: ATP; autophosphorylation; phosphotransferase; tyrosine-specific protein kinase
 F:129-188/Domain: immunoglobulin homology <IMM>
 F:589-932/Domain: protein kinase homology <KIM>
 F:597-605/Region: protein kinase ATP-binding motif

Query Match 24.0%; Score 1263.5; DB 1; Length 978;
 Best Local Similarity 33.8%; Pred. No. 2.3e-54;
 Matches 334; Conservative 171; Mismatches 361; Indels 121; Gaps 37;

35 CVLHKNNDSSVGVKSSYPMVSESPEDLGCALRPOSSGTVYEAADVDSASITLQVL 94
 12 CVLLVLLRGQTGTSQPSAGPEPSP-----STQPAQS-----ELIVEAGDTIRLT 57
 95 VDAPGNISCLWVFKHSLNQCQPHFDLQNRGVWSVILKMTPE-TQAGEYLLFIQSEATNYT 153
 58 CTDPAFVK-WTFE-----LLDVRIENKQSEWREKAEATHGKYTCVSSGLRSSI 107
 154 ILFTVSIIRN-TLTYLRRPYFRKMNQDALVCISVPEPVEVVLCDSGESCKESPA 212
 108 YVF---VRDPAVFLVGLPLFGK-EDNDALVRCLPTDPQ-VSNYSLEICDKGLPTDLKF 162
 213 V-----VKKEKVLHFLFGTDIRCCARNELGRECTRLFTID-----LNQTFQTLPLQ 260
 163 VPNPKAGITIKNVKRAYHRLC---IRCAQREGKWMRSDFTLKVRAAIKAIPIWVSPET 219
 261 --FLKVGPELWIRCKAVHNVHGFGLTW-----ELENKALIEEGNVFEMS---TYSTNRMTIR 311
 220 SHLLKGGDTFTVICTKIDVSTSDVMWIKLNPQSKAQVKRNSWQGDVNYERQETLT- 278
 312 ILFAFVSSVARNDTGYTSCSSKHPQSALVT---IVGKGFN---ATNSSEDIYDOVE 365
 279 -----ISSARVNDSGVFMVYANNFTGSANVTTLKVKVEKGFNIFPVKNIT--VFVTDGE 331
 366 EFCFSVRFKAYPO-IRCTWTFSRKSPFCQKGLD-----NGYSISKFCNKH-----QP 413

332 NVDLVVEFEAYPKPEHQWIYMRN--PTNRGEDIYKSDNQSNIRYVNELELRLTKGTEG 389
 414 GYVIFHAENDDAQFTQWFTLNIRRKQVLA--EASASQASCFSDGYPLRPSWTWKCKSDKS 471
 390 GTYTFVLVNSDSVASVTFDYYVNTKPEILTYDLKMMGRLOQVAAAGFPFPIIDYWFCTGAE 449
 472 PNCETEBITE-GVNNRKANRKFQGVWSSSTLNNSEAIKGLFKVCCAYNSLGTSCETILLN 530
 450 QRCITVPVPVDVQIQNASVSPFGKLVQSSIDSSVFRHNGTVECKASNAVGS--SAFFN 507
 531 SPGPPPP-----IQDN-----ISFYATIGVCLLFIIVLTLCHTKYKKKFRYESQL 576
 508 ----FAFKGNSKEIQIOPHTLFTPLIGFVVVTA--LMGIIVMVL--AYKYLQKPMYEVQW 559
 577 QMV-QVTGSSDNEFYFVVDREYEDLKWEFPRENLEFGVLGSGAGKVMNATAYGISKT 635
 560 KVEEING---NNYVYIDPTQPYDHKWEFPNRNLSFGKTLGAGAFKVVVEATAYGLIKS 616
 636 GVSIOVAVKMLKEKADSSREALMSELKMTQLGSHENIVNLLGACTLSGPIVLIIFYCC 695
 617 DAAMTVAVKMLPESAHLTREALMSELKVLVSLGNHNVNLLGACTVGGPTLVITEYCC 676
 696 YGDLNLNLSKREKHFRTWTEIFKEHNFSPYTFQSHPNMSPGSRH-VQIHPSDDOISG 754
 677 YGDLNLFRLRKDSF--IFSKQEBEQADALYKNLLHSKSSCDSSNEYMDMKPGVSYVWP 734
 755 LHGNSFHS---EDEIYENQKLEEEEDLNLTFLFCFAYOVAKGMEFLKPSKCVHRD 811
 735 TKTKERSARIYSIERDVTTPAIMEDEL-ALDELLSFSYQVAKGMFLASKNCTHRD 793
 812 LAARNVLVTHGVKVICDFGLARDIMSDSNVYVRGNARLPVKWMAPELFGIYTIKSDV 871
 794 LAARNILLTHGRITKICDFGLARDIMSDSNVYVRGNARLPVKWMAPELFGIYTIKSDV 853
 872 WSYGILLWEIFSLGVNYPGIPVDANFYKLIQNGFQMDQPFYATEEIIYIMQSCWAFDSR 931
 854 WSYGIFLWELFSLGSSPYGMPVDSKPYKWKIEGFRMLSEHAPAAWYEVKTCWDADPL 913
 932 KPSFNLTSFLGCLADAEAEAMQNV 958
 914 KRPTFKQVQLIEKQISDSTKHYSNL 940

RESULT 7
 TVHUKT
 protein-tyrosine kinase (EC 2.7.1.112), receptor type kit precursor - human
 N:Alternate names: mast/stem cell growth factor receptor; tyrosine kinase receptor c-kit
 C:Species: Homo sapiens (man)
 C:Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 21-Jul-2000
 C:Accession: S01426; PC1015; A41815; B41815; I37948; I56954; I54336
 R:Yarden, Y.; Kuang, W.J.; Yang-Feng, T.; Coussens, L.; Munemitsu, S.; Dull, T.J.; Chen
 EMBO J. 6, 3341-3351, 1987
 A:Title: Human proto-oncogene c-kit: a new cell surface receptor tyrosine kinase for an
 A:Reference number: S01426; MUID:88111521; PMID:2448137
 A:Accession: S01426
 A:Molecule type: mRNA
 A:Residues: 1-976 <YAR>
 A:Cross-references: GB:X06182; NID:g34084; PIDN:CAA29548.1; PID:g34085
 R:Hu, W.X.; Cornu, F.; Andre, C.; Galibert, F.
 Chinese Biochem. J. 7, 618-629, 1991
 A:Title: Nucleotide sequence of two neighbouring fragments of human c-kit proto-oncogene
 A:Reference number: PC1015
 A:Accession: PC1015
 A:Molecule type: DNA
 A:Residues: 412-713 <HUM>
 A:Note: article in Chinese with English abstract
 R:Spritz, R.A.; Giebel, L.B.; Holmes, S.A.
 Am. J. Hum. Genet. 50, 261-269, 1992
 A:Title: Dominant negative and loss of function mutations of the c-kit (mast/stem cell ;
 A:Reference number: A41815; MUID:92133600; PMID:1370874
 A:Accession: A41815
 A:Molecule type: DNA
 A:Residues: 579-583, 'L', 585-589 <SPR>

A;Cross-references: GB:S78839; NID:G244084; PIDN:AB21234.1; PID:G244085
A;Note: sequence extracted from NCBI backbone (NCBIN:78839, NCBI:P:78842)
A;Note: disease-related mutant from patient with piebaldism
A;Accession: B41815
A;Molecule type: DNA
A;Residues: 637-641, 'SPELPW' <SP2>
A;Cross-references: GB:S78843; NID:G244086; PIDN:AB21235.1; PID:G244087
A;Note: sequence extracted from NCBI backbone (NCBIN:78843, NCBI:P:78844)
A;Note: disease-related mutant from patient with piebaldism
A;Accession: C41815
A;Molecule type: DNA
A;Residues: 556-560, 'GDKWK' <SP3>
A;Cross-references: GB:S78845; NID:G244088; PIDN:AB21236.1; PID:G244089
A;Note: sequence extracted from NCBI backbone (NCBIN:78845, NCBI:P:78846)
A;Note: disease-related mutant from patient with piebaldism
R;Giebel, L.B.; Strunk, K.M.; Holmes, S.A.; Spritz, R.A.
Oncogene 7, 2207-2217, 1992
A;Title: Organization and nucleotide sequence of the human KIT (mast/stem cell growth factor receptor) gene
A;Reference number: 137948; MUID:93064697; PMID:1279499
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-976 <RES>
A;Cross-references: EMBL:X69301; NID:G34089; PIDN:CAA9159.1; PID:G82586
A;Note: an alternative splice form omitting residues 510-513 is described
R;Yamamoto, K.; Tojo, A.; Aoki, N.; Shibuya, M.
Jpn. J. Cancer Res. 84, 1136-1144, 1993
A;Title: Characterization of the promoter region of the human c-kit proto-oncogene.
A;Reference number: 156954; MUID:94103107; PMID:7506248
A;Accession: 156954
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-22 <RES>
A;Cross-references: GB:S67773; NID:G459358; PIDN:AA29529.1; PID:G459359
R;Spritz, R.A.; Holmes, S.A.; Berg, S.Z.; Nordlund, J.J.; Fukai, K.
Hum. Mol. Genet. 2, 1499-1500, 1993
A;Title: A recurrent deletion in the KIT (mast/stem cell growth factor receptor) proto-oncogene
A;Reference number: 154336; MUID:94061059; PMID:7694728
A;Accession: 154336
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 242-250 <RES>
A;Cross-references: GB:S67686; NID:G460545; PIDN:AD13996.1; PID:G4261696
A;Gene: GDB:KIT
A;Cross-references: GDB:120117; OMIM:164920
A;Map position: 4q12-q12
A;Introns: 23/1; 113/1; 207/1; 252/3; 309/1; 372/2; 411/1; 449/2; 514/1; 549/3; 592/1; 607/1
A;Note: defects in this gene may result in piebaldism
C;Function:
A;Description: catalyzes the phosphorylation of a peptidyl tyrosine residue by ATP
C;Superfamily: macrophage colony-stimulating factor 1 receptor; immunoglobulin homology;
C;Keywords: alternative splicing; ATP; autophosphorylation; glycoprotein; kinase-related
protein; tyrosine-specific protein kinase
F;1-976/Product: protein-tyrosine kinase kit precursor, long form #status predicted <MAT>
F;1-509,514-976/Product: protein-tyrosine kinase kit precursor, short form #status predicted <SIG>
F;1-22/Domain: signal sequence #status predicted <SIG>
F;23-976/Product: protein-tyrosine kinase kit #status predicted <MAT>
F;23-520/Domain: extracellular #status predicted <EXT>
F;51-99/Domain: immunoglobulin homology <IMM1>
F;129-188/Domain: immunoglobulin homology <IMM2>
F;226-292/Domain: immunoglobulin homology <IMM3>
F;328-394/Domain: immunoglobulin homology <IMM4>
F;423-493/Domain: immunoglobulin homology <IMM5>
F;521-543/Domain: transmembrane #status predicted <TM>
F;544-976/Domain: intracellular #status predicted <INT>
F;587-931/Domain: protein kinase homology <KIN>
F;595-603/Region: protein kinase ATP-binding motif
F;58-97, 136-186, 233-290, 428-491/Diulfide bonds: #status predicted
F;130, 145, 283, 293, 300, 320, 352, 367, 463, 486/Binding site: carbohydrate (Asn) (covalent) #
F;623, 640, 792/Active site: Lys, Glu, Asp #status predicted
F;797, 810/Binding site: magnesium (Asn, Asp) #status predicted

Query Match 23.7%; Score 1251; DB 1; Length 976;
Best Local Similarity 32.7%; Pred. No. 9.2e-54;
Matches 324; Conservative 166; Mismatches 352; Indels 148; Gaps 34;
QY 47 VGKSSVPMVSESPEDLGCALRQSSQGTVEAAAVEVDVSASITLOVLVDAPGNISCLMW 106
DB 20 VQTSSQSPSPG-EPSPPSIHFGKSDLI-----VAVGEIRLLCTDTPGVK-WT 67
QY 107 FKHSLLCNCPHFLQNRGVVSMVLKMTETQAGEYLLFIQSEATN---YTLFTVSIKRT 163
DB 68 FE-----ILDETENKQEWIT-EKAEATNCTKCTNKHLSNS 106
QY 164 L-----LYTLRRPYFRKMENQDALVCISSEVPPIVEWVLCDSQGSCKEE----- 209
DB 107 IYFVRDPAKFLVDRSLYGK-EDNDTLVRCPLTDP-VTNYSLKGCQGRPLKPLRPIP 164
QY 210 ---SPAVKKEEKLHELFGTDIRCCARNELGRECTRLFTIDL---NOTPQTTLPO--L 260
DB 165 DPVAGIMIKSVKAYHRLC---LHCSVDQBGKSVLSEKFLKVRPAKAVPVSVSKASY 221
QY 261 FLKVGEPFLWIRCAVHVNHGFLTWELN---KALEEGNYFEMSTYSTRNTRMIRILFAFV 317
DB 222 LLAREGEFTVCTIKOVSSSVYTWKRENSQTKLOEKYNSMHHGDFNVERQAT-----LTI 277
QY 318 SSVARNDTGYTCCSSSKHPQSALVT---IVGKGFINA-----TNSSEDEIYOYE 365
DB 278 SSARVNDSGVFMFCYANNTFGSANVTTLLEVVDKGFIFPMINTVTFVNDGENDVL----- 333
QY 366 EFCFSYRFKAYPO-IRCTWTFSRKSF--PCEQKGLDNGYSISKFCNHK-----OPGE 415
DB 334 ----IVEYEAFFKPEHQMIYMRNRTFDKWEDVPKSESNIRYVSELHLTRLKGTGGT 389
QY 416 YIFHAENDDAQTKMFTLNIRRPQVILA--EASASQASCFSDGVPPLPSWTWKKCSKSPN 473
DB 390 YTLVNSNDVNAIAFNYYVNTKPEILTYDLVNGMLQCVAAAGEPETIDWYFCPTQEQ 449
QY 474 CTBEITE-GVNNRKANKKVGQWSSSTLANSBAIKGFLVCCAYNSLGTSCETILNSP 532
DB 450 CSASVLPVDVQTLNSSGPPFKLVQSSIDSSAFKHNGTVCKRAYNDVGKT--SAYFN-- 505
QY 533 GPPFTODN-----ISFYATIGVCLLFTVLTLLCHYKQKQFRYESOLOM 578
DB 506 --FAPKNNKEQTHPHTLFTPLLIGFVIVAGMCCIIVMLT---YKLOKPMTEVOMKV 559
QY 579 V-QVTSSDNEYFYVDPREYEDLKWEFFRENLEFGVLGSGAPGKVMNATAYGISTGV 637
DB 560 VEING---NNYVYIDPTQLPYDHKEFFPNRLSFGKTLGAGAGKVVETAYGLIKSDA 616
QY 638 SIQAVVQLKEKADSSERREALMSELKMMTOLGSHENIVNLLGACTLSGPIYLIPEYCCY 697
DB 617 AMTVAVKMLKPSAHLTEREALMSELKVLVSLGNHMINVNLGACTIGTPTLVITEYCCY 676
QY 698 DILNYSLSKEKEHRTWTE-----TFKEHNFSEYPTFQSHPNSSM---PGRREVOIHPD 748
DB 677 DLLNLFRRKEDSTICQKQEDHAEALYKNNLLHSEKSCSDSTNBYMDKPGVSTV-VPTK 735
QY 749 SDIISGLHGSFSEDEIEYENKRLBEERDLNVLTPEDLLCFAYQVAKGMEFELEKSCV 808
DB 736 ADKRSVRIGSY-----IERDVTPIAIMEDEL-ALDLELLLSFSYQVAKGNAFLASKNCI 789
QY 809 HRDLAARNLVTHGKVKIKICDFGLARDIMSDSNVYVRGNARLPVKWMAPELSEFEGITYIK 868
DB 790 HRDLAARNLLTHGRIKICDFGLARDIKNDSNVYVRGNARLPVKWMAPELSEFEGITYIK 849
QY 869 SDVWSYGIILWEIPLSGVNPYGPIDANFYKLIQNGFKMDQPPYATEEYIINQSCWAF 928
DB 850 SDVWSYGIILWEIPLSGVNPYGPIDANFYKLIQNGFKMDQPPYATEEYIINQSCWAF 928
QY 929 DSKRSESPNLTSLFLGCOLADAEEAMQNV 958
DB 910 DPLKRPPTFKOIVOLIEKQISESTNHHYSNL 939

RESULT 8
 151703
 C-kit-related kinase 1 (XKrk1) - African clawed frog
 C:Species: Xenopus laevis (African clawed frog)
 C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 18-Jun-1999
 C:Accession: 151703
 R:Baker, C.V.; Sharpe, C.R.; Torpey, N.P.; Heasman, J.; Wylie, C.C.
 Mech. Dev. 50, 217-228, 1995
 A:Title: A Xenopus c-kit-related receptor tyrosine kinase expressed in migrating stem cell
 A:Reference number: 151703; PMID:95344996; PMID:7619732
 A:Accession: 151703
 A:Status: preliminary; translated from GE/EMBL/DDBJ
 A:Molecule type: mRNA
 A:Residues: 1-954 <BAK>
 A:CROSS-references: EMBL:Z48770; NID:G763033; PIDN:CAA8688.1; PID:G763034
 C:Superfamily: macrophage colony-stimulating factor 1 receptor; immunoglobulin homology
 C:Keywords: ATP
 F:575-915/Domain: protein kinase homology <KIN>
 F:583-591/Region: protein kinase ATP-binding motif

Query Match 23.2%; Score 1226; DB 2; Length 954;
 Best Local Similarity 33.1%; Pred. No. 1.5e-52;
 Matches 305; Conservative 162; Mismatches 322; Indels 132; Gaps 30;

QY 107 FXHSLNCPHFDLQNRGVVS-----MVLKMTQAGEYLLFIQSEATNYTILFTVSI 160
 DB 52 FOKSGLMKPR-DLKRPLNNSETDQFVLLKADLRHIGRYCTNTQENTSV--SLVF 108
 QY 161 RNTLLYLRPRFRKMNODALVCSIESPEPIVEWVLCDSQGESCKE-----ESPA 212
 DB 109 KDPARPFLOIPFDVTEGADTCMCFPTDPMIAIEKCD--GSLPENFTFTDIEAGI 166
 QY 213 VVKKEKVLHFGTDIRCARNELGR-ECTRLFTDLNQTPTTLPOLFL-----KV 264
 DB 167 TIKTQALFDSQV-----VCSGNKSGTVKSSFTSHVVPVKR-KVPTVFLSKSRQLVKT 220
 QY 265 GSEPLIRCKAVVNHGFLTW-ELENKALBEGYFEMSTYSTNRTRMIRILFAFVSSVARN 323
 DB 221 GEPFEVTCVAVLDVFTVKAQMLDKGVTKQANFRSNFVSNLTLS-----DGVPSV 274
 QY 324 DTCYTCSSSKHPSQ---SALVTIVGKGFINAT-NSSEYEDIDQEEFCFSVRFAKAYPO- 378
 DB 275 ESRFTQENALGQVNAFTLVDVGVNLTVENTIISVAGNVLVVKVVIDAYPH 334
 QY 379 IRCTWTFSSKSPFCEQGLDNGYSISK-----FCNHKH-----QPGYIIFHAENDDA 425
 DB 335 DGVVNTYFNETL---LNTSDHYVATKDEGNRYVSELHLRLKGTKEGVYTYTNSDD 390
 QY 426 QTKMFTLNIRKPKQVLAASASOAS--CFSDGYPLPSTWTKKCDKSPNCTBEITEGVW 483
 DB 391 DASVSFNIQVTRPELIIAERTSEGLQCVATGFPVPAIQWYFCPSGEQRCTDYPPLSPV 450
 QY 484 NRK--ANRKFVGQWSSSTLNKSEALKGLVAKCAVNSLGTSCETILLNSPGPFPE-IOD 540
 DB 451 NEKFTQENSLGRIVVESIIDVNDLKNGTQCVASNEVESAYSV-----FSFAIKE 502
 QY 541 N-----ISFYATIGVCLLFTVLVTLILCHIKKQFYESQLQWY-QVTGSSDNE 588
 DB 503 KLRTHTLFTPLLIGFTAAAGLMCIAVAVL---MYKYLQPKVEIQWVVEEING---NN 555
 QY 589 YFVDFREYEDLKWEFFPRENLEFGKVLGSGAPGKWNATAYISIKTVSIVQAVQMLKE 648
 DB 556 YVYIDPTQLPYDNKWEFFRDLRCFGILGAGAFGVKEATAYGLLKEDSRLLTAVAKMLXP 615
 QY 649 KADSSERBALMELKQMTOLGSHENTVNLGLACTLSGPYLVLFYCCYGDLLNLYLSKRE 708
 DB 616 SAHSTEREALMELKVLVSLGHKKNLVNLLGACTVGGPILVITEYCCYGDLLNLYLRKRD 675
 QY 709 KPHRTWTEIFKEHNFYSFYTFQSHPS-----SMPGSREVQIHPDSQOI----- 752
 DB 676 SF-----ICPKFEDNSEAALYKKNLNTDMCGECSMEYIDMKPAVSVVVPT 721
 QY 753 ---SGLHNSFHSEDELEYENQKRLBEEDLNVLTFEDLLCFAYQVAKGMEFLFKSCVHR 810

DB 722 KTDKRSGSGFGQDV-----SVSIPEDDL-ALDTEDLNFSYQVAGQMNFLASKNCIHR 775
 QY 811 DLAAARNVLVTHGKVKVCDPGLARDIMSDSNVYVVRGNARLPIVKWMAPESLFEGIYTIKSD 870
 DB 776 DLAAARNILLTHGRITKICDFGLARDIRNDSNVYVVGKGNARLPIVKWMAPESFHCVTYTFSD 835
 QY 871 VNSYGTLLWEISLGNVPYGPVDFANFYKLLONGFKMDQPPFVATBEIYIIMQSCWAFDS 930
 DB 836 VNSYGTLLWEISLGNVPYGPVDFANFYKLLONGFKMDQPPFVATBEIYIIMQSCWAFDS 895
 QY 931 RKRPSPFNLTSLFGCOLADAE 951
 DB 896 LKRPFKQIVQWVEQQLSDSK 916

RESULT 9
 TVCTMD
 N:Contains: protein-tyrosine kinase (EC 2.7.1.112) csflr/fms
 C:Species: Felis silvestris catus (domestic cat)
 C:Date: 31-Dec-1989 #sequence_revision 31-Dec-1989 #text_change 13-Jun-1997
 C:Accession: A31636
 R:Woolford, J.; McAuliffe, A.; Rohrschneider, L.R.
 Cell 55, 965-977, 1988
 A:Title: Activation of the feline c-fms proto-oncogene: multiple alterations are required
 A:Reference number: A31636; PMID:89077553; PMID:2849512
 A:Molecule type: mRNA
 A:Residues: 1-980 <WOO>
 A:CROSS-references: EMBL:X03663
 C:Genetics:

C:Superfamily: macrophage colony-stimulating factor 1 receptor; immunoglobulin homology;
 C:Keywords: ATP; autophosphorylation; glycoprotein; kinase-related transforming protein;
 fic protein kinase

F:1-23/Domain: signal sequence #status predicted <SIG>
 F:24-380/Product: macrophage colony-stimulating factor 1 receptor #status predicted <SIG>
 F:24-509/Domain: extracellular #status predicted <EXT>
 F:35-86/Domain: immunoglobulin homology <IMM1>
 F:120-179/Domain: immunoglobulin homology <IMM2>
 F:217-280/Domain: immunoglobulin homology <IMM3>
 F:316-381/Domain: immunoglobulin homology <IMM4>
 F:410-484/Domain: immunoglobulin homology <IMM5>
 F:510-534/Domain: transmembrane #status predicted <TM>
 F:535-980/Domain: intracellular #status predicted <INT>
 F:577-915/Domain: protein kinase homology <KIN>
 F:585-593/Region: protein kinase ATP-binding motif
 F:42-42,127-177,224-278,417-482/Disulfide bonds: #status predicted
 F:45,73,94,153,275,302,335,410,477,490/Binding site: carbohydrate (Asn) (covalent) #stat
 F:613,630,776/Active site: Lys, Glu, Asp #status predicted
 F:781,794/Binding site: magnesium (Asn, Asp) #status predicted

Query Match 23.2%; Score 1225.5; DB 1; Length 980;
 Best Local Similarity 33.5%; Pred. No. 1.6e-52;
 Matches 321; Conservative 168; Mismatches 321; Indels 149; Gaps 34;

QY 83 VDVSASITLQVLVDAPGNISCLW---VFKHSLNCP-----HFDLQNRGVSMVIL 131
 DB 32 VEPGTVTLRCV-----GNGSVWDGPISEHWNLDLDPSSILTTNNATQNTGYHCT-- 85
 QY 132 KMTTQAGEYL-LFTQSEATNYTIL---FTVSIRNTLLYTLRPPYFRKMNODALV-CI 185
 DB 86 EPGNFGGNATILYVKDPAKPKVLAQEVTV-----LEGDALLPCL 128
 QY 186 SSVSEPEPIVE--WVLCDSQGESCKE-----SP---AVVKEEKVLHFGTDIRCARN 235
 DB 129 ---LTDPALEAGVSLVRVRGPELVLRQNTYNSFSGHGTTHKAKFIENHYV---QCSARV 181
 QY 236 ELGRECTRL-----FTIDLNQTPTTL-PQLFLKV-GEPLWIRCKAVVNHGFLTWELE 288
 DB 182 D-GRVTSGIWLKVKQKQISGPATITLPEALVRIQGEAAQIVCSASNIDVNFV----- 235

Qy 836 IMSDSYVVRGNARLPVKWAPESLFEIGYTIKSDVMSYGILLWEIFSLGVNPPGPVD 895
 Db 801 IMDSNYIVKGNARLPVKWAPESLFCVYTVQSDVMSYGILLWEIFSLGVNPPGPILVN 860
 Qy 896 ANFYKLQNGKMDQPFYATIEIYIMQSCWAFSRKPSFNPNTSFLGCOLADAEAMY 955
 Db 861 SKFYKLVDGYQMAQAPAPAKNIYSIMQACWALEPRTPTFOQICSLLOKQ---AQE--- 914
 Qy 956 QNVGRV---SECHTYQNR--RPF 975
 Db 915 ---DRVENVNTLPSSSSSRLRPW 936

RESULT 11
 TVHMD
 N:Macrophage colony-stimulating factor 1 receptor precursor - human
 C:Species: Homo sapiens (man)
 C>Date: 28-Dec-1987 #sequence revision 31-Dec-1993 #text change 16-Jul-1999
 C:Accession: S08123; A24533; I56672; I57648; I59083; I52772
 R:Hampe, A.; Shamoon, B.W.; Gobet, M.; Sherr, C.J.; Galibert, F.
 Oncogene Res. 4, 9-17, 1989
 A:Title: Nucleotide sequence and structural organization of the human FMS proto-oncogene
 A:Reference number: S08123; MUID:89239490; PMID:2524025
 A:Accession: S08123
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-972 <HAM>
 A:Cross-references: GB:U63963; EMBL:X14720; NID:g1915975; PIDN:AA851696.1; PID:g1915976
 A:Note: this sequence was submitted to the EMBL Data Library, March 1989
 R:Cousens, L.; Van Beveren, C.; Smith, D.; Chen, E.; Mitchell, R.L.; Isacke, C.M.; Verm
 Nature 320, 277-280, 1986
 A:Title: Structural alteration of viral homologue of receptor proto-oncogene fms at carb
 A:Reference number: A24533; MUID:86175013; PMID:2421165
 A:Accession: A24533
 A:Molecule type: mRNA
 A:Residues: 1-53, 'A', 55-972 <COU>
 A:Cross-references: GB:J03149
 A:Note: the authors translated the codon GCA for residue 54 as Pro
 R:Wheeler, E.F.; Roussel, M.F.; Hampe, A.; Walker, M.H.; Fried, V.A.; Look, A.T.; Retten
 J. Virol. 59, 224-233, 1986
 A:Title: The amino-terminal domain of the v-fms oncogene product includes a functional e
 sequences.
 A:Reference number: I56672; MUID:86281820; PMID:3525854
 A:Accession: I56672
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-16 <RES>
 A:Cross-references: GB:M14002; NID:g182676; PIDN:AAA35849.1; PID:g553292
 R:Visavader, J.; Verma, I.M.
 Mol. Cell. Biol. 9, 1336-1341, 1989
 A:Title: Differential transcription of exon 1 of the human c-fms gene in placental troph
 A:Reference number: I57648; MUID:89261741; PMID:2524648
 A:Accession: I57648
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-16 <RE2>
 A:Cross-references: GB:M25786; NID:g349454; PIDN:AAA58421.1; PID:g553224
 R:Browning, P.J.; Bunn, H.F.; Cline, A.; Shuman, M.; Nienhuis, A.W.
 Proc. Natl. Acad. Sci. U.S.A. 83, 7800-7804, 1986
 A:Title: Replacement of COOH-terminal truncation of v-fms with c-fms sequences markedl
 A:Reference number: I59083; MUID:87017034; PMID:3532121
 A:Accession: I59083
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 874-972 <RE3>
 A:Cross-references: GB:M14193; NID:g182521; PIDN:AAA35834.1; PID:g182522
 R:Nienhuis, A.W.; Bunn, H.F.; Turner, P.H.; Gopal, T.V.; Nash, W.G.; O'Brien, S.
 Cell 42, 421-428, 1985
 A:Title: Expression of the human c-fms proto-oncogene in hematopoietic cells and its del
 A:Reference number: I52772; MUID:85282599; PMID:4028159
 A:Accession: I52772
 A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA
 A:Residues: 244-295 <RE4>
 A:Cross-references: GB:M11067; NID:g182674; PIDN:AAA35848.1; PID:g442423
 C:Genetics:
 A:Gene: GDB:CSF1R; FMS
 A:Cross-references: GDB:120600; OMIM:164770
 A:Map position: 5q33.2-5q33.3
 A:Introns: 17/1; 103/1; 198/1; 243/3; 297/1; 361/2; 400/1; 440/2; 504/1; 542/3; 585/1;
 C:Superfamily: macrophage colony-stimulating factor 1 receptor; immunoglobulin homology;
 C:Keywords: ATP; autophosphorylation; glycoprotein; kinase-related transforming protei
 fic protein kinase
 F:1-23/Domain: signal sequence #status predicted <SIG>
 F:24-972/Product: macrophage colony-stimulating factor 1 receptor #status predicted <W
 F:24-512/Domain: extracellular #status predicted <EXT>
 F:35-86/Domain: immunoglobulin homology <IMM1>
 F:120-179/Domain: immunoglobulin homology <IMM2>
 F:217-280/Domain: immunoglobulin homology <IMM3>
 F:316-383/Domain: immunoglobulin homology <IMM4>
 F:413-487/Domain: immunoglobulin homology <IMM5>
 F:513-537/Domain: transmembrane #status predicted <TM>
 F:538-972/Domain: intracellular #status predicted <INT>
 F:580-917/Domain: protein kinase homology <KIN>
 F:588-596/Region: protein kinase ATP-binding motif
 F:42-84, 127-177, 224-278, 419-485/Diulfide bonds: #status predicted
 F:45, 73, 153, 240, 275, 302, 335, 353, 412, 428, 480/Binding site: carboxydrate (Asn) (covalent)
 F:616, 633, 778/Active site: Lys, Glu, Asp #status predicted
 F:783, 796/Binding site: magnesium (Asn, Asp) #status predicted

Query Match 23.1%; Score 1216.5; DB 1; Length 972;
 Best Local Similarity 32.8%; Pred. No. 4.4e-52;
 Matches 325; Conservative 156; Mismatches 316; Indels 195; Gaps 33;

Qy 75 YVEAAVEVDV--SASITLQVLVAPGNISCLWFKHSLNCQPHFDLQNRGVNVLK 132
 Db 22 VIEPSVPELVVPGATVTLRCV---GNGSVEMDGPSS---PHWTLYSDGSSILSTN 72
 Qy 133 MTE-----TQAGEYL-----LFIQSEATNTLTFTVSIINTLLYTLRRYFRKM 176
 Db 73 NATQNTGYTRCTPGDPLGSAIHLVKDPAAPNNVL-----AOEVVVF----- 118
 Qy 177 ENQDALV-CISESPEPIVEVWLCDSCQESCKEESPAAVKKEEK-----VLHE--- 223
 Db 119 EDQALLPCL---LTDPVLE-----AGVSLVRVGRPLMRHTNYSFSPWHGFTTHRAKF 169
 Qy 224 LFGTDIRCARNEIGRECTRLFTIDLNOTPTLP-----OLFLKVGPELWIRCK 273
 Db 170 IQSDYQCSALMG-GR---KVMISIRLKVQKVPGLPPALTLVPAELVIRGEAAQIVCS 225
 Qy 274 AVHNHGELTWELNKALEBGNFEMSTYTRNTRMIRILFAFVSSVARNDTGYTCSSS 333
 Db 226 ASSVDVNFDFVLOHNNTKLA---IPQSDFFHNNRYQ-KVLTLNLDQVDFQHAGNYSCVAS 281
 Qy 334 ---KHPSQSALVTIVGKGFNATNSED---YEIDQVEECFVSFRKAYPOIR-CTWTF 385
 Db 282 NVQSKH-STSMFFRVESAYLNL---SSEQNLIOEVTVGEGNLKVMVEAYPLGQGFNWTY 338
 Qy 386 SRKSPPCQKGLDNGYSISKFCNKHQP-----LGPESDQPEPLKANATTKDYRHTLTSLPRKSEAGYSF 418
 Db 339 -----LGPESDQPEPLKANATTKDYRHTLTSLPRKSEAGYSF 381
 Qy 419 HAENDDAQTKMTINTIRKQVLAESAASQAS-----CFSDGYPLPSTWTKCKSKSPNC 474
 Db 382 LARNPGGWRALTFTLTLYRPPESVVIWTFINGSGTLLCAASGPPQPNVTWLCQSGHTDRC 441
 Qy 475 TEEITEGVWN---RKANRKYFGQWVSSSTINMSEAIKGFVLCVCCAYNSLGTSCETILLN 530
 Db 442 DEAQVLQWDDPYDEVUSQEPFHKVTVOSLTVTETLEHNTQTECAHNSVSGGSWAFPI 501
 Qy 531 SPGPPFPFTQDNISFYATIGVC---LLFIWVLTLLICHKYKQFYRESQIQWVQVGTSSDN 587
 Db 502 SAGATHPPDFELFTPPVVVACMSIMALLLLLLLLLYKKQKPKYQVRWKIIE--SYEGN 559
 Qy 588 BYFYVDREYEDLKWEPFRENLEFGKVLGSAFGKVNATAYGISKTGVSIVQAVRMKJ 647

Db 560 SYTFIDPTQLPYNEKWEFPRNNLQFGKTLGAGAFKVVETAFGLGKEDAVLKAVKMLK 619
 Qy 648 EKADSSEREALMSLKMWTOLGSHENIVNLLGACTLSGPIYILIFYCCYGLNLYLSRKR 707
 Db 620 STAHADKEALMSLKMWTOLGSHENIVNLLGACTLSGPIYILIFYCCYGLNLYLSRKR 679
 Qy 708 E-----KFHTWTETFEKHNFSFYPTFQSHPNSSMPGSRVQIHP 747
 Db 680 EAMLGPSLSPGQDPGGVDYKNIHLKRYVRDGSFS-----SQGVDTYV 724
 Qy 748 DSDQISGLHGSFHEDEIEYENQKRLKEEDLNVLTPEDLLCFAYQVAKGMERLEFKSC 807
 Db 725 EMRPVSTSSNSF-SEQDLKEDGRLP-----LRDLHFSQVAGKAFVLSKNC 774
 Qy 808 VHRDLAARNVLVTHGKVKIKDFGLARDIMSDSNVVRGNARLPVKWMAPELSFEGIYTI 867
 Db 775 IHRDVAARNVLVTHGKVKIKDFGLARDIMSDSNVVRGNARLPVKWMAPELSFEGIYTI 834
 Qy 868 KSDVMSYGILLWEIFSLGVNVPYGPVDANFYKILQNGFKMDQPPYATEEIIYIMQSCWA 927
 Db 835 QSDVMSYGILLWEIFSLGVNVPYGPVDANFYKILQNGFKMDQPPYATEEIIYIMQSCWA 894
 Qy 928 FDSRKRPFPNLTSLGCOL-ADAEAMQNV 958
 Db 895 LEPTHRPTFOQICSLPQAOEDRRERYTNL 926

RESULT 12

JN0677
 protein-tyrosine kinase (EC 2.7.1.112), receptor type kit precursor - chicken
 N;Alternate names: tyrosine kinase receptor kit
 A;Title: Cloning and expression of the chicken c-kit proto-oncogene.
 A;Reference number: JN0677; MUID:93292995; PMID:7685729
 C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 16-Jun-2000
 C;Accession: JN0677
 R;Sasaki, E.; Okamura, H.; Chikamune, T.; Kanai, Y.; Watanabe, M.; Naito, M.; Sakurai, M.
 Gene 128, 257-261, 1993
 A;Cross-references: DBJ:D13225; NID:g303532; PID:BAA02506.1; PID:g303533
 A;Experimental source: brain
 C;Superfamily: macrophage colony-stimulating factor 1 receptor; immunoglobulin homology;
 C;Keywords: ATP; autophosphorylation; glycoprotein; kinase-related transforming protein;
 protein kinase
 F;1-24/Domain: signal sequence #status predicted <SIG>
 F;25-960/Product: tyrosine kinase receptor #status predicted <WAT>
 F;314-380/Domain: immunoglobulin homology <IMW>
 F;573-916/Domain: protein kinase homology <KIN>
 F;581-589/Region: protein kinase ATP-binding motif
 F;76,135,149,269,286,306,318,338,356,453,469/Binding site: carbohydrate (Asn) (covalent)

Query Match 23.0%; Score 1215; DB 1; Length 960;
 Best Local Similarity 32.3%; Pred. No. 5.1e-52;
 Matches 314; Conservative 165; Mismatches 370; Indels 124; Gaps 34;
 Qy 66 ALRPOSSGTVYEAAYVDVDSITQLVLDVAPGNISCLVVKHSS-----LNCQPHFD 119
 Db 18 SLIPAGSVPHBESSLVNKGELKCNKEGP-----VTWNFQNSDPSAKTRISNEKEWH 73
 Qy 120 LQNRGV--VSMVILKMTQAGELYLFIQSEATNYTLFTSVIRNLTLLYLRPFRKME 177
 Db 74 TKNATIRDIGRECKSGKSVNSFYVFKDP-----NVLFV---DSLIV-----GKED 119
 Qy 178 NODALVCSISEVPEP-IVBWLICDSOGESCKEESPAVVKKE-----KVLHFLFGTDIRCC 232
 Db 120 SDILLVC---PLTDPDLNFTLRKCDGKLPKXMTPIPNPQKGIILIKVQRFKGYCYQL 176
 Qy 233 AR-NELGRECTRLFTIDLQNP-QITLPLQ-----FLKVGELPLWIRCKAVHNVHFGI 283
 Db 177 AKHNGVEKISEHF---LNVPRVHKALPVITLKSVELLKEGEEFEVTCITTDVSSVKA 233

Qy 284 TWELNKALBEGNFEMFSTYSTNETMIRILFAPVSSVARNDTGYVTCSSSKHP--SQSAL 341
 Db 234 SWISYKSAIVTSKSRNLDGYGERK-----LTLNRSVGUNDSGEFTC-QAENPFKTKNAT 288
 Qy 342 VTI--VGKGFN--ATNSSDEYIDQYEBFCFSVRFKAYPQIR-CTWTFGRKSPFCBQK 395
 Db 289 VTLKALAGFVRLPATWNTTIDINAGQNG--LIVEYEAYPKPEEVWYNNETL--QN 343
 Qy 396 GLDNGYSIKFCNKKH-----OPEYIIPHAENDDAQFTYKMTFLNIRRPQVLA 443
 Db 344 SSDHVVKPFTVGNNSYSELHLRLKGTGGIYTFVFNSSDASSSVTFNVVYKTKPEILT 403
 Qy 444 EASASQ--ASCFSOGYPLPSMTWKCKSDKSPNCTEITEGVNWRKANRKFQWVSSSTL 501
 Db 404 LDMGLNDILQCVATGFPAPTYYVYPCGTQRCLODSDPTISPMDDVKV-----YTNSSVP 457
 Qy 502 NMSEAIKGLFKCCAYNSLGTSCETILLNSPGPPF-----IQDNISFYA-----TIGV 550
 Db 458 SFERILVESTVNASMFKSTGTCCEASNGDKSSVFFNFAIKEQIRHTLPTPLLIAGV 517
 Qy 551 CLLFIVWLTLICHKKYKQFYESQLQMV-QVTGSSDNEYFYVDREYEDLKWFFPREN 609
 Db 518 AAGLMCIIVMLVILYLOKPKYEQWKVVEING--NNYVYIDPTQLPYDHKWFPPNR 574
 Qy 610 LEFGKVLGSGAFGKVMNATAYGISTGVSIQVAVKMLKEKADSSREALMSLKMWTQLG 669
 Db 575 LSPGKTLGAGAFKGVVEATAYGLFKSDAAMTVAVKMLKPSAHLTEREALMSLKVLSYL 634
 Qy 670 SHENIVNLLGACTLSGPIYILIFYCCYGLNLYLSRKR-----HRTWTEIFKEHNFSF 725
 Db 635 NHINIVNLLGACTLSGPIYILIFYCCYGLNLYLSRKR-----HRTWTEIFKEHNFSF 692
 Qy 726 YPTFQSHPNSS-----SMPGSRVQIHPDSDQISGLHGSFHEDEIEYENQKRLKEE 778
 Db 693 --LHQAEPTADAVNEMDMKPGVSVAVPPKADKKRPVSGSYTDQD-----VTLSMLEDD 746
 Qy 779 DLNVLTFFDLICFAYQVAKGMERLEFKSCVHRDLAARNVLVTHGKVKICDFGLARDIM 838
 Db 747 --LALDVEDLISFSYQVAKGMSFLAKNCIHRDLAARNVLVTHGKVKICDFGLARDIM 804
 Qy 839 DSNVYVRGNARLPVKWMAPELSFEGIYTIKSDVMSYGILLWEIFSLGVNVPYGPVDANF 898
 Db 805 DSNVYVRGNARLPVKWMAPELSFEGIYTIKSDVMSYGILLWEIFSLGVNVPYGPVDANF 864
 Qy 899 YKLIQNGFMQDPFYATEEIIYIMQSCWAFSRKRPFPNLTSLFGLQADAEAMQNV 958
 Db 865 YKMIKEGYMFSPECSPEMYDIMKSCWDADPLQRTFTKQIVQLIEQQLSDNAPRYAN- 923
 Qy 959 DGRVSECPHTYQN 971
 Db 924 ---FSTPPSTQCN 933

RESULT 13

TJ30816
 macrophage colony-stimulating factor receptor - Japanese pufferfish
 C;Species: Fugu rubripes (Japanese pufferfish)
 C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 02-Jun-2000
 C;Accession: TJ30816
 R;How, G.F.; Venkatesh, B.; Brenner, S.
 Genome Res. 6, 1185-1191, 1996
 A;Title: Conserved linkage between the pufferfish (Fugu rubripes) and human genes for p.
 A;Reference number: Z20882; MUID:97129405; PMID:8973913
 A;Accession: TJ30816
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-975 <HOW>
 C;Cross-references: EMBL:U63926; NID:g1752706; PID:g1752708; PID:AA60063.1
 C;Superfamily: macrophage colony-stimulating factor 1 receptor; immunoglobulin homology;
 Query Match 23.0%; Score 1213; DB 2; Length 975;
 Best Local Similarity 32.4%; Pred. No. 6.5e-52;

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2003, 07:14:08 ; Search time 19.0096 Seconds
(without alignments)
2456.527 Million cell updates/sec

Title: US-09-919-408A-4
Perfect score: 5274
Sequence: 1 MPALARDAGTVPLLVFSAM.....PFSREMDLGLSPQAQVDS 993

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5274	100.0	993	1 FLT3 HUMAN	P36888 homo sapien
2	4429.5	84.0	992	1 FLT3 MOUSE	Q00342 mus musculus
3	1295.5	24.6	978	1 KIT CARPI	Q28317 capra hircu
4	1286	24.4	977	1 KIT BOVIN	P43481 bos taurus
5	1266	24.0	975	1 KIT MOUSE	P05532 mus musculus
6	1251	23.7	976	1 KIT HUMAN	P10721 homo sapien
7	1247.5	23.7	975	1 KIT CANFA	O97799 canis fami
8	1225.5	23.2	980	1 KFS FELCA	P13369 felis silve
9	1224.5	23.2	978	1 KIT FELCA	Q28889 felis silve
10	1224	23.2	978	1 KFS FSVMD	P00545 feline sarc
11	1216.5	23.1	972	1 KFS HUMAN	P07333 homo sapien
12	1215	23.0	960	1 KIT CHICK	Q08156 gallus gall
13	1204	22.8	977	1 KFS MOUSE	P09581 mus musculus
14	1194	22.6	978	1 KFS RAT	Q00495 rattus norv
15	1176	22.3	1088	1 FGDS RAT	P20786 rattus norv
16	1169	22.2	1089	1 PGDS MOUSE	P26618 mus musculus
17	1166	22.1	1089	1 PGDS HUMAN	P16234 homo sapien
18	1128	21.4	1087	1 PGDS XENLA	P26619 xenopus lae
19	1105.5	21.0	1098	1 PGR MOUSE	P05622 mus musculus
20	1086.5	20.6	1106	1 PGR HUMAN	P09619 homo sapien
21	990	18.8	1338	1 VGR1 HUMAN	P17948 h vascular
22	989.5	18.2	370	1 KIT FSVHZ	P04048 feline sarc
23	959.5	18.2	1333	1 VGR1 MOUSE	P35969 mus musculus
24	956	18.1	1336	1 VGR1 RAT	P53767 rattus norv
25	951.5	18.0	1356	1 VGR2 HUMAN	P35968 homo sapien
26	935	17.7	1348	1 VGR2 COTJA	P52583 coturnix co
27	923	17.5	1367	1 VGR2 MOUSE	P35918 mus musculus
28	917.5	17.4	1343	1 VGR2 RAT	O08775 rattus norv
29	907.5	17.2	1363	1 VGR3 MOUSE	P35917 mus musculus
30	904.5	17.2	1298	1 VGR3 HUMAN	P35916 homo sapien
31	874.5	16.6	823	1 CEK3 CHICK	P18461 gallus gall
32	849	16.1	821	1 FGR2 HUMAN	P21802 mus sapien
33	845	16.0	813	1 FGR2 XENLA	Q00364 xenopus lae

ALIGNMENTS

RESULT 1

FLT3_HUMAN STANDARD; PRT; 993 AA.

AC P36888; Q13414;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE FL cytokine receptor precursor (EC 2.7.1.112) (Tyrosine-protein kinase receptor FLT3) (Stem cell tyrosine kinase 1) (STK-1) (CD135 antigen).
GN FLT3 OR STK1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP MEDLINE=94119906; PubMed=7507245;
RA Small D., Levenstein M., Kim E., Carow C., Amin S., Rockwell P., Witte L., Burrow C., Ratajczak M.Z., Gewirtz A.M., Civin C.I.;
RT "STK-1, the human homolog of Flk-2/Flt-3, is selectively expressed in CD34+ human bone marrow cells and is involved in the proliferation of early progenitor/stem cells.";
RT Proc. Natl. Acad. Sci. U.S.A. 91:459-463(1994).
RL [2]
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=93357464; PubMed=8394751;
RA Rosnet O., Schiff C., Pebusque M.J., Marchetto S., Tonnelle C., Toiron Y., Birg F., Birnbaum D.;
RT "Human FLT3/FLK2 gene: cDNA cloning and expression in hematopoietic cells.";
RL Blood 82:1110-1119(1993).
RN [3]
RP SEQUENCE OF 783-942 FROM N.A.
RX MEDLINE=91169547; PubMed=2004790;
RA Rosnet O., Mattei M.-G., Marchetto S., Birnbaum D.;
RT "Isolation and chromosomal localization of a novel FMS-like tyrosine kinase gene.";
RL Genomics 9:380-385(1991).
CC -!- FUNCTION: RECEPTOR FOR THE FL CYTOKINE. HAS A TYROSINE-PROTEIN KINASE ACTIVITY.
CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein tyrosine phosphate.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- TISSUE SPECIFICITY: BONE MARROW CELLS.
CC -!- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-PROTEIN KINASES.
CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
CC -!- DATABASE: NAME=PROW; NOTE=CD guide CD135 entry;
CC -!- WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd135.htm".
CC -!- DATABASE: NAME=Atlas Genet. Cytogenet. Oncol. Haematol.;
CC -!- WWW="http://www.infobiogen.fr/services/chronocancer/Genes/FLT3ID144.html".

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; U02687; AA18947.1; -;
 CC EMBL; 226582; CAAB1393.1; -;
 CC EMBL; L36162; AAA35487.1; -;
 CC PIR; A39061; A39061.
 CC HSP; P11362; 1FGK.
 CC Genew; HGNC:3765; FLT3.
 CC MIM; 136351; -;
 CC GO; GO:0005887; C-integral to plasma membrane; TAS.
 CC GO; GO:0005021; F-vascular endothelial growth factor receptor; TAS.
 CC GO; GO:0008284; P-positive regulation of cell proliferation; TAS.
 CC GO; GO:0007169; P-transmembrane receptor protein tyrosine kin. . .; TAS.
 CC InterPro; IPR007110; IG-like.
 CC InterPro; IPR003599; IG.
 CC InterPro; IPR003006; IG_MHC.
 CC InterPro; IPR000719; Prot_kinase.
 CC InterPro; IPR001824; TKKinaseII.
 CC InterPro; IPR001245; Tyr_pkinase.
 CC Pfam; PF00047; ig; 1.
 CC Pfam; PF00069; pkinase; 1.
 CC ProDom; PD000001; Prot_kinase; 2.
 CC SMART; SM00409; IG; 1.
 CC SMART; SM00219; TyKc; 1.
 CC PROSITE; PS00835; IG LIKE; 1.
 CC PROSITE; PS00107; PROTEIN KINASE ATP; 1.
 CC PROSITE; PS00011; PROTEIN KINASE DOM; 1.
 CC PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
 CC PROSITE; PS00240; RECEPTOR_TYR_KIN_III; 1.
 CC Signal; Transferase; Tyrosine-protein kinase; Receptor; Transmembrane;
 KW Glycoprotein; Phosphorylation; ATP-binding; Immunoglobulin domain.
 FT SIGNAL 1 26
 FT CHAIN 27 993
 FT DOMAIN 27 543
 FT TRANSMEM 544 563
 FT DOMAIN 564 593
 FT DOMAIN 593 624
 FT DOMAIN 624 644
 FT NP_BIND 644 644
 FT BINDING 644 644
 FT ACT_SITE 811 811
 FT CARBOHYD 43 43
 FT CARBOHYD 100 100
 FT CARBOHYD 151 151
 FT CARBOHYD 306 306
 FT CARBOHYD 323 323
 FT CARBOHYD 351 351
 FT CARBOHYD 354 354
 FT CARBOHYD 473 473
 FT CARBOHYD 502 502
 FT CARBOHYD 541 541
 FT CONFLICT 8 8
 FT CONFLICT 10 11
 FT CONFLICT 78 78
 FT CONFLICT 227 227
 FT CONFLICT 346 346
 FT CONFLICT 940 940
 FT SEQUENCE 993 AA; 112804 MW; 16790124B02F6BBF CRC64;

Query Match 100.0%; Score 5274; DB 1; Length 993;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 993; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MPALARDAGTVPLLVVFSAMIFGTITNQDLPVTKCVLINHKNNDSSVGKSSSYPMVSESP 60
 DB 1 MPALARDAGTVPLLVVFSAMIFGTITNQDLPVTKCVLINHKNNDSSVGKSSSYPMVSESP 60
 QY 61 EDLGCALRQSSGTVYAAAVEVDVVSASITLQVLDPAGNISCWLFKHSLLNCQPHFDL 120
 DB 61 EDLGCALRQSSGTVYAAAVEVDVVSASITLQVLDPAGNISCWLFKHSLLNCQPHFDL 120

QY 121 QNRGVSVWILKMTQAGEYLLFTIOSEATNTYILFTVSRNTLLYTLRRPYRKMNQD 180
 DB 121 QNRGVSVWILKMTQAGEYLLFTIOSEATNTYILFTVSRNTLLYTLRRPYRKMNQD 180
 QY 181 ALVCISESVPEPIVWVLCDSOGESCKERSPAVVKKEKVLHFLGTDIRCCARNELGRE 240
 DB 181 ALVCISESVPEPIVWVLCDSOGESCKERSPAVVKKEKVLHFLGTDIRCCARNELGRE 240
 QY 241 CTRLFIDLNQPTTLPOLFLKVGEPILWIRCKAVHNVHGFGLTWELKALEEGNYFEM 300
 DB 241 CTRLFIDLNQPTTLPOLFLKVGEPILWIRCKAVHNVHGFGLTWELKALEEGNYFEM 300
 QY 301 STYSTNRTMIRILFAFVSSVARNDTGYYTCSSSKHPSQSALVTIVGKGFNATNSSDEYE 360
 DB 301 STYSTNRTMIRILFAFVSSVARNDTGYYTCSSSKHPSQSALVTIVGKGFNATNSSDEYE 360
 QY 361 IDQYEEFCFSVRPKAYPQIRCTWTFPSRKFPCEBQKGLDNGYSISKFNHKKHQGEYIFHA 420
 DB 361 IDQYEEFCFSVRPKAYPQIRCTWTFPSRKFPCEBQKGLDNGYSISKFNHKKHQGEYIFHA 420
 QY 421 ENDDAQFTKMTINIRRKQVLAESAASQSCFSDGYPLPSWTWKCKSDKSPNCTEITE 480
 DB 421 ENDDAQFTKMTINIRRKQVLAESAASQSCFSDGYPLPSWTWKCKSDKSPNCTEITE 480
 QY 481 GWNKRNKRVFGQWVSSSTLNMSEAIKGLVKCCAYNSLGTSCETILLNSPGPPFIQD 540
 DB 481 GWNKRNKRVFGQWVSSSTLNMSEAIKGLVKCCAYNSLGTSCETILLNSPGPPFIQD 540
 QY 541 NISPVATIGVCLLFIWVLLTLLCHKVKYKOFYVESQIQWQVTVGSSDNEYVYDFREYED 600
 DB 541 NISPVATIGVCLLFIWVLLTLLCHKVKYKOFYVESQIQWQVTVGSSDNEYVYDFREYED 600
 QY 601 LKWEFPRENLEFGKVLGSGAFGKVMNATAYGISTGVSIOVAVKMLKEKADSSEREALMS 660
 DB 601 LKWEFPRENLEFGKVLGSGAFGKVMNATAYGISTGVSIOVAVKMLKEKADSSEREALMS 660
 QY 661 ELKMMTQLGSHENIVNLLGACTLSGPIYLIFFEYCCYGDLLNLYRSKREKPHRTWTETPKE 720
 DB 661 ELKMMTQLGSHENIVNLLGACTLSGPIYLIFFEYCCYGDLLNLYRSKREKPHRTWTETPKE 720
 QY 721 HNPFSYPTFQSHPNSSMPGSRVQIHPSDQISGLHNSFHSDELEYENKLEEDL 780
 DB 721 HNPFSYPTFQSHPNSSMPGSRVQIHPSDQISGLHNSFHSDELEYENKLEEDL 780
 QY 781 NVLTPEDLFCFAYQVAKGMFELEFKSCVHRDLAARNVLVTHGKVKICDFGLARDIMSDS 840
 DB 781 NVLTPEDLFCFAYQVAKGMFELEFKSCVHRDLAARNVLVTHGKVKICDFGLARDIMSDS 840
 QY 841 NYVVRGNARLPVKWMAPESLFEGITYIKSDVWSYGILLWEIFSLGVNYPGIPVDANFYK 900
 DB 841 NYVVRGNARLPVKWMAPESLFEGITYIKSDVWSYGILLWEIFSLGVNYPGIPVDANFYK 900
 QY 901 LIQNGFKMDOPFYATEEIIYIMQSWAFDSRKPSPFNLTSLFLGCOLADAEAMYNQVDG 960
 DB 901 LIQNGFKMDOPFYATEEIIYIMQSWAFDSRKPSPFNLTSLFLGCOLADAEAMYNQVDG 960
 QY 961 RVSECPHTYQNRPPSRREMDLGLLSPQAQVEDS 993
 DB 961 RVSECPHTYQNRPPSRREMDLGLLSPQAQVEDS 993

RESULT 2
 FLT3_MOUSE
 ID FLT3_MOUSE STANDARD; PRT; 992 AA.
 AC Q00342;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE FL cytokine receptor precursor (EC 2.7.1.112) (Tyrosine-protein kinase
 receptor flk-2) (Fetal liver kinase 2) (Tyrosine-protein kinase FLT3).
 GN FLT3 OR FLT-3 OR FLK-2.
 OS Mus musculus (Mouse).


```

Db 661 SELAMTHLGHONIIVNLLGACTLUSGVPYLIFCYCCGDLNLYRSKREPHRTWTIEFK 720
Qy 720 EHNFSYPTQSHPNSSMPGSRVQIHPDSDQISGLHSGNFSHSEDEIYENQKEL--EEE 777
Db 721 EHNFSYPTQAHSSNSMPGSRVQIHPDSDQISGLHSGNFSHSEDEIYENQKELABEE 780
Qy 778 EDNLVITFEDLLCFAYQAVKGMFLEPKSCVHRDLAARNVLVTHGKVVVKICDFGLARDIM 837
Db 781 EDNLVITFEDLLCFAYQAVKGMFLEPKSCVHRDLAARNVLVTHGKVVVKICDFGLARDIL 840
Qy 838 SDSNYVVRGNARLVPKVMAPESLPEGIYTKSDVMSYGILLWEIYFSLGVNYPGIPVDAN 897
Db 841 SDSYVVRGNARLVPKVMAPESLPEGIYTKSDVMSYGILLWEIYFSLGVNYPGIPVDAN 900
Qy 898 FYKLQNGFMDOPFVATEIYIMQSCWAFDRKRPSPNLTSLFCGLQCLADAEAEAMYQN 957
Db 901 FYKLQSGFKMEQFPYATEGIYFWQSCWAFDRKRPSPNLTSLFCGLQCLADAEAEAC-- 957
Qy 958 VGRVSECPHYQNRPRFSEMDLGLLSPQAQVE 991
Db 958 ----IRTSIHLPKQAAPQQRG-GLRAQSPQRQVK 986

RESULT 3
KIT_CAPIH STANDARD; PRT; 978 AA.
AC Q28317;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Mast/stem cell growth factor receptor precursor (EC 2.7.1.112) (SCFR)
DE (Proto-oncogene tyrosine-protein kinase Kit) (c-kit).
GN KIT.
OS Capra hircus (Goat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Capra.
OX NCBI_TaxID=9925;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Shiba; TISSUE=Cerebellum;
RX MEDLINE=97342548; PubMed=919245;
RA Tanaka S., Yanagisawa N., Tojo H., Kim Y.-J., Tsujimura T.,
RA Kitamura Y., Sawasaki T., Tachi C.;
RT "Molecular cloning of cDNA encoding the c-kit receptor of Shiba goats
RT and a novel alanine insertion specific to goats and sheep in the
RT kinase insert region.";
RL Biochim. Biophys. Acta 1352:151-155(1997).
CC -!- FUNCTION: THIS IS THE RECEPTOR FOR STEM CELL FACTOR (MAST CELL
CC GROWTH FACTOR). IT HAS A TYROSINE-PROTEIN KINASE ACTIVITY.
CC BINDING OF THE LIGANDS LEADS TO THE AUTOPHOSPHORYLATION KIT AND
CC ITS ASSOCIATION WITH SUBSTRATES SUCH AS PHOSPHATIDYLINOSITOL
CC 3-KINASE (PI3K).
CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC tyrosine phosphate.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- SIMILARITY: BELONGS TO THE CSP-1/PDGF RECEPTOR FAMILY OF TYROSINE-
CC PROTEIN KINASES.
CC -!- SIMILARITY: Contains 5 immunoglobulin-like C2-type domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; D45168; BAA08116.1; --
DR HSSP; P11362; 1PGK.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003599; Ig.

```

```

DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR001824; RTKinaseIII.
DR InterPro; IPR001245; Tyr_kinase.
DR Pfam; PF000047; ig; 1.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 2.
DR SMART; SM00409; IG; 3.
DR SMART; SM00219; TYRK; 1.
DR PROSITE; PS00835; IG_LIKE; 2.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE; PS00240; RECEPTOR_TYR_KIN_III; 1.
KW Proto-oncogene; Tyrosine-protein kinase; Receptor; Transmembrane;
KW Transferrase; Glycoprotein; Phosphorylation; ATP-binding; Signal;
KW Immunoglobulin domain.
FT SIGNAL 1 22
FT CHAIN 23 978
FT DOMAIN 23 521
FT TRANSMEM 522 544
FT DOMAIN 545 578
FT DOMAIN 590 939
FT NP_BIND 596 604
FT BINDING 624 624
FT ACT_SITE 794 794
FT MOD_RES 825 825
FT CARBOHYD 94 94
FT CARBOHYD 130 130
FT CARBOHYD 145 145
FT CARBOHYD 284 284
FT CARBOHYD 294 294
FT CARBOHYD 301 301
FT CARBOHYD 321 321
FT CARBOHYD 353 353
FT CARBOHYD 368 368
FT CARBOHYD 401 401
FT CARBOHYD 464 464
FT CARBOHYD 487 487
SQ SEQUENCE 978 AA; 109722 MW; CA4D663F98205CA9 CRC64;

Query Match 24.8%; Score 1295.5; DB 1; Length 978;
Best Local Similarity 33.3%; Pred. No. 4.2e-81;
Matches 335; Conservative 169; Mismatches 339; Indels 163; Gaps 35;

Qy 47 VGHSSYVPVSVSPEDLGC-ALRPQSGTYVEAAVEVDVSASITLOVLDAPGNISCLW 105
Db 20 VQTGSSQPSV--SPGELSFSIHPAKS-----ELIVSGVDEIRLLCTDPGFVK--W 66
Qy 106 VFKHSLNCPHDLQNRGVSMVLKMTQAGEYLLFIQSEAT---NYTILFTVSI RN 162
Db 67 TFE-----ILQQLSEKTNPEWIT-EKAEATNGTCTNKGGLSS 105
Qy 163 TL-----LYTLRRPFRKMNQDALVCISEVPEPIVEWVLCDSQGCKEESPAV 213
Db 106 SIYVFVRDEPKFLIDLPLVYKKEEN-DTLVRCPLDPE-VTNYSLTGCEGKPLPKDLTFV 163
Qy 214 -----VKKEEKLHELFGTDIRC--CARNELGRE-CTRLFTID----LNQTPQTLTP 258
Db 164 ADPKAGITIRNVKREYHRL-----CLHCSANQKGSMLSKFTLVRAAKAVPVVSVS 217
Qy 259 Q--LFLKVGELWIRCKAVVNHGFLGTWELENKALEEGNYPEMSTYSTNRTWIRILFAP 316
Db 218 KTSYLLREGEERFAVTCIKDVSSVDSVMIKENSQOSKQATKKNWSHQGDFSVLRQERLT 277
Qy 317 VSSVARNDTGYTCCSSKHPQSALVT---IVGKGFINA-----TNSSDEYDIQY 364
Db 278 ISSARVNDGVMFCYANNFTFGSANVTTLEVDVKGFINFPMMNTTVFVNDGENVDL--- 334
Qy 365 EEFCSVRFKAYPQ-----IRCTWTSRKSFPCEQKGLDNGYSIKPCNHKH----- 411
Db 335 -----VVEYEAVPKPEHQRWIMYMNRTSTDKWDYPKSE-----NESNIRYVVELHLTRLK 384

```

QY 412 --QPEYIFHAENDDAQFTMETLNRKQVLAESA--SQASCSDGCVLPSTWKEC 467
 DB 385 QTEGTYTHFVNSDVSNSVTVNNVNTPEILTHDLVNGMLQCAAGFPPEIDWYFC 444
 QY 468 SDKSPNCTEEI--TEGVNKKANKRVGOWSSSTLNMSEAIKGLVKCCAYNSLGTSCRT 526
 DB 445 PCTEORCSVPVGVQVQIONSSVSPGKLVVSTIDDSFTFKNGITVECRAYNDVGKSSAS 504
 QY 527 ILLNSPGPPFTQDN-----ISFVATIGVCLLFVVLVTLILCHIKYKQFQY 572
 DB 505 F-----NFAFGKNNKEQIHAHTLFTPLLIGFVIAAGLMCIFWMLT-----YKYLQKPMY 554
 QY 573 ESQLOMV-QVTSSDNEYFVDFREYDOLKWEFFRENLERGKVLGSGAGFKVMATAYG 631
 DB 555 EYQWKVVEING---NNVYVYIDPTQLFYDHPKWEFFPRNLSFGKTLGAGAFKQVVEATAYG 611
 QY 632 ISKTGVSIOVAVKMLKEKADSEREALMSELKMTOLGSHENIVNLLGACTLSGPTLYLIF 691
 DB 612 LIKSDAAMTVAVKMLKPSAHLTEREALMSELKVLVSLGNHMIIVNLLGACTIGGPTLVIT 671
 QY 692 EYCCYGDLLNLYLSRREKPHRTWTE-----IPKEHNFYFPTFQSHPNSSM---PGSRE 742
 DB 672 EYCCYGDLLNLFRRKRDSPICQOEBAEVALYKVLNLLHSSKSSCNDSTNEYMDMKPGVSY 731
 QY 743 VOIHPDSDOISGLHNSFHSDEIEYENOKRLEEDLNLVTFEDLLCFAYQVAKGMEFL 802
 DB 732 VVPTKAADRRSARIGSY-----IERDVTFAIMEDDEL-ALDIEDLSFSYQVAKGMFL 785
 QY 803 EFKSCVHRDLAARNVLVTHGKVKICDFGLARDIMSDSNVVRGNARLPVKWAPESLFE 862
 DB 786 ASKNCHIRDLAARNLLTHGRITKICDFGLARDIKNSNVVKGARLPVKWAPESIFN 845
 QY 863 GYTIKSDVWSYGILLWEIFSLGVNPGIPVDANFYKLIQNGKMDQPFYATEEIIYM 922
 DB 846 CVYTSEDVWSYGIFLWELFSLGSSYPGMPVDSKEYKIKGFRMLSPBEHAPAEYDIM 905
 QY 923 QSCWAFSDKRSPFNLTSLFGLCOLADAEAMVQNVGDRVSEC-PH 967
 DB 906 KTCMDADPLKPTFKQIVOLIBKQISESTNHIYSN-----LANGSPH 947
 RESULT 4
 KIT_BOVIN
 ID KIT_BOVIN STANDARD; PRT; 977 AA.
 AC P43481;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DE Mast/stem cell growth factor receptor precursor (EC 2.7.1.112) (SCFR)
 DE (Proto-oncogene tyrosine-protein kinase kit) (c-kit).
 GN KIT.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OX NCBI TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94215924; PubMed=7512939;
 RA Kubota T., Hikono H., Sakaki E., Sakurai M.;
 RT "Sequence of a bovine c-kit proto-oncogene cDNA";
 RL Gene 141:305-306(1994).
 CC -!- FUNCTION: THIS IS THE RECEPTOR FOR STEM CELL FACTOR (MAST CELL
 CC GROWTH FACTOR). IT HAS A TYROSINE-PROTEIN KINASE ACTIVITY.
 CC BINDING OF THE LIGANDS LEADS TO THE AUTOPHOSPHORYLATION KIT AND
 CC ITS ASSOCIATION WITH SUBSTRATES SUCH AS PHOSPHATIDYLINOSITOL
 CC 3-KINASE (PI3K).
 CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
 CC tyrosine phosphate.
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -!- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-
 CC PROTEIN KINASES.
 CC -!- SIMILARITY: Contains 5 immunoglobulin-like C2-type domains.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; D16680; BAA04084.1; -;
 CC HSSP; P11362; 1FGK.
 DR InterPro; IPR007110; IG-like.
 DR InterPro; IPR003599; IG.
 DR InterPro; IPR003006; IG_MHC.
 DR InterPro; IPR000719; Prot_kinase.
 DR InterPro; IPR001824; RTKaseIII.
 DR InterPro; IPR001245; Tyr_kinase.
 DR Pfam; PF00047; ig; 1.
 DR Pfam; PF00069; pkinase; 1.
 DR ProDom; PD000001; Prot_kinase; 2.
 DR SMART; SM00409; IG; 3.
 DR SMART; SM00219; TyRK; 1.
 DR PROSITE; PS00835; IG_LIKE; 2.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00240; RECEPTOR_TYR_KIN_III; 1.
 DR PROSITE; PS00240; RECEPTOR_TYR_KIN_III; 1.
 KW Proto-oncogene; Tyrosine-protein kinase; Receptor; Transmembrane;
 KW Transferase; Glycoprotein; Phosphorylation; ATP-binding; Signal;
 KW Immunoglobulin domain.
 FT SIGNAL 1 22
 FT CHAIN 23 977
 FT DOMAIN 23 521
 FT TRANSMEM 522 544
 FT DOMAIN 545 977
 FT DOMAIN 590 938
 FT NP_BIND 596 604
 FT BINDING 624 624
 FT ACT_SITE 793 793
 FT MOD_RES 824 824
 FT CARBOHYD 94 94
 FT CARBOHYD 130 130
 FT CARBOHYD 145 145
 FT CARBOHYD 284 284
 FT CARBOHYD 294 294
 FT CARBOHYD 301 301
 FT CARBOHYD 321 321
 FT CARBOHYD 353 353
 FT CARBOHYD 368 368
 FT CARBOHYD 401 401
 FT CARBOHYD 464 464
 FT CARBOHYD 487 487
 SQ SEQUENCE 977 AA; 109685 MW; 4B2719050883B7EF CRC64;
 Query Match 24.4%; Score 1286; DB 1; Length 977;
 Best Local Similarity 33.2%; Pred. No. 1.9e-80;
 Matches 332; Conservative 172; Mismatches 344; Indels 152; Gaps 36;
 QY 47 VCKSSSYNWSVSPEDLGC-ALRPQSSGVTYEAAVEVDVSASITQLVLDAGNISCLW 105
 DB 20 VQTGSSQPSV--SPGELSIPSHPAKS-----ELIVSGDEIRLLCTDGFVK--W 66
 QY 106 VFKHSSLNCQPHFDLQNRGVSMVLKMTQTQAGEVLLFIOSEAT---NYTILFTVSIRN 162
 DB 67 TFE-----ILQLSEKNTPWIT-EKAEATNGVTCNKGGLSS 105
 QY 163 TL-----LYTLRRPYFRKMNQDALVCISESVPEPIVEWVLCDSCSKESPAV 213
 DB 106 SIYVVRDPEKFLDIDLPLYGKEEN-DTLVRCLTDPE-VTNVSLGCGKPLPKDLTFV 163
 QY 214 -----VKKEKVLHFLGFTDIRC--CARNELGRE-CTRLFTID-----LNOTPQTLLP 258
 DB 164 ADPKAGITIRNVKREYHRL-----CLHSANQRGKSMLSKRFLLKVRRAIKAVPVVSVS 217

DR InterPro; IPR001245; Tyr_pkinase.
 DR Pfam; PF00047; ig; 1.
 DR Pfam; PF00069; pkinase; 1.
 DR ProDom; PD000001; Prot_kinase; 2.
 DR SMART; SM00408; IGC2; 1.
 DR SMART; SM00219; TyrKc; 1.
 DR PROSITE; PS00835; IG LIKE; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
 DR PROSITE; PS00240; RECEPTOR_TYR_KIN_III; 1.
 KW Proto-oncogene; Tyrosine-protein kinase; Receptor; Transmembrane;
 KW Transferrase; Glycoprotein; Phosphorylation; ATP-binding; Signal;
 KW Immunoglobulin domain; Disease mutation.
 FT SIGNAL 1 22
 FT CHAIN 23 975
 FT DOMAIN 23 975 MAST/STEM CELL GROWTH FACTOR RECEPTOR.
 FT TRANSMEM 23 519 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 520 542 POTENTIAL.
 FT DOMAIN 543 975 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 588 935 PROTEIN KINASE.
 FT NP_BIND 594 602 ATP (BY SIMILARITY).
 FT BINDING 622 622 ATP (BY SIMILARITY).
 FT ACT_SITE 790 790 BY SIMILARITY.
 FT MOD_RES 821 821 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
 FT CARBOHYD 146 146 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 296 296 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 303 303 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 323 323 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 355 355 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 370 370 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 466 466 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 489 489 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT VARIANT 1 12 MRGARGWDLIC -> MAVAVFPPLPQQ (IN TRUNCATED FORM).
 FT VARIANT 13 785 MISSING (IN TRUNCATED FORM).
 FT VARIANT 582 582 E -> K (IN W37 SPOTTING).
 FT VARIANT 660 660 T -> M (IN W37 SPOTTING).
 FT VARIANT 790 790 D -> N (IN W42 SPOTTING).
 FT VARIANT 831 831 V -> M (IN W42 SPOTTING).
 SQ SEQUENCE 975 AA; 109001 MW; BAB5CA4D9AF9CD2A CRC64;

Query Match 24.0%; Score 1266; DB 1; Length 975;
 Best Local Similarity 33.2%; Pred. No. 4.4e-79;
 Matches 336; Conservative 173; Mismatches 330; Indels 172; Gaps 41;

QY 35 CVLINHKNDSSVSGKSSYPMVSPEDLGCALRPQSGTGYEAAAVEVDVSASITLQVL 94
 DB 12 CVLLVLLRGQTATQPSASPGSPSP-----SIHPAQ-----ELIVEAGDTL--- 54
 QY 95 VDAPGNISCL-----WVFHSSLNCQPHPLQNGVSVWILKWTQAGEYLLFTQSE 148
 DB 55 -----SLTCDPDPFVRWTFK-----TYFN-----EMVENKQNEWIQ-EKAE 89
 QY 149 ATNYTLFTVSIKNTL-----LYTLRRPYFRKMNODALVCISSEYPIPEIVE 195
 DB 90 ATR-TGTYTCSNLSGLTSSIVYFVRDPAKLVGLPLFGK-EDSDALVRCPLTDPQ-VSN 146
 QY 196 WVLCDSCGESCCKERSPAV-----VKKEKVLHFLFGTDIRCCARNELGREGCTRL--- 244
 DB 147 YSLIECDGKSLPTDLTPVPNPKAGITIKNVKAYHRLC-----VRCQAQ-----RDGTWLHSD 199
 QY 245 -FTIDLNOT-----PQTLPLQ-----FLKVGEPILWIRCKAVHNVHGFGLTW-----E 286
 DB 200 KFTLKVREAIKAIKIPVSVVPETSHLLKXGDTPTVVTCTIKDVSTSVNSMWLKNPQPOHIAQ 259
 QY 287 LENKALBEGNFENSTYNTNRTMLRILFAFVSSVARNDTGYYTCSKSPQSALVT--- 343
 DB 260 VKHNSWHRGDF-----NVEROETLT-----ISSARVDDSGVFMFYANNTPFGSANVTTLK 309
 QY 344 IVKGKGFINATN-SEDEYIDQYEFCEFSVRKAYPQ-IRCTWTFSRKSFPCEQKGLD--- 398
 DB 310 VVEKGFINISPVKNITTVFDGENDVLDVVEEAYPKPEHQQWIYNNRT--SANKGKYVK 367

QY 399 -NGYSISKFCNNHKH-----QPGEYIFHAENDDAQFTKMTFLNIRKPOVLA--EASAS 448
 DB 368 SDNKSNIYVYNQLRLTLKGTGEGTYFLVSNDSASASVTENVYNTKPEILTYDRLNG 427
 QY 449 QASCFSDGYPLPSWTKKCDKSPNCTEETITE-GVMNRKANRKFVGQWSSSTLMSBAI 507
 DB 428 MLQCVAEGFPEPTIDWYFCTGAEQRCITPSPVQVQVQVQVSPFGKLVQVSSIDSSVFR 487
 QY 508 KGFLVKCCAYNSLGTSCETILLNSPGPPF---TODN-----ISFVATTGVCLLFIWV 557
 DB 488 HNGTVECKASNDVGKS--SAPFN---FAFKEQIQAHTLFTPLLIGFVVAAGAMGIIVMV 541
 QY 558 LTLILCHIKKQFYESOLOMV-OVTGSSDNEYFVDPREYEDLKWKFFPRENLEBFGKVL 616
 DB 542 LT-----YKLOKPNYEVQMKVVEING---NNYVIIDTQPLPDYHKWFFPRNLSFGKTL 594
 QY 617 GSGAFGKVMNATAYGISKTGVSIQVAVKMLKEKADSSEREALMSLKMXTOLGSHENIVN 676
 DB 595 GAGAFGKVEATAYGLIKSDAAMTAVKMLKPSAHLTEREALMSLKVLSVLGNHMINVN 654
 QY 677 LLGACTLSGPIYLIFFYCCYCGDILLNLYLSRKEP-----HRTWTEIFKEHNFSPYTFQ 730
 DB 655 LLGACTYGGPTLVITYCYCCYCGDILLNLYLSRKEP-----HRTWTEIFKEHNFSPYTFQ 714
 QY 731 SHPNSSM---PGSREVOIHPDSDOISGLHNSPHSEDEIEYENQKRLSEEDLNVLTFED 787
 DB 715 S-SNEYMDMPGVSYV-VPTKIDKRRSARDISY-----IERDVTPTAIMEDEL-ALDLD 766
 QY 788 LLFCAYQVAKMEFLFKSCVHRDLAARNVLVTHGKVKVICDFGLARDIMSDSNVVRGN 847
 DB 767 LLSFSYQVAKAMAFASKNCIHRDLAARNILLTHGRITKICDFGLARDIRDSNVVVKGN 826
 QY 848 ARLPVKWAPESLPEGITYTKSDVWSYGIILLWEIPSLGVNPPYGPIDVDFYKLIQNGFK 907
 DB 827 ARLPVKWAPESIPSCVYTFESDVWSYGIIFLWELFSLGSSPYGMPVDSKFKYKMEKGR 886
 QY 908 MDOPPYATEEIIYIMQSCWAFDSRPSFNLTLFGLCOLADABEAMQNV 958
 DB 887 MVSEHAPAEYMDVMKTCWDADPLKRPTEFKQVQLIBKQISDSTKHIYSNL 937

RESULT 6
 KIT_HUMAN
 ID KIT_HUMAN STANDARD; PRT; 976 AA.
 AC P10721;
 DT 01-JUL-1989 (Rel. 11, Created)
 DT 01-JUL-1989 (Rel. 11, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Mast/stem cell growth factor receptor precursor (EC 2.7.1.112) (SCFR)
 DE (Proto-oncogene tyrosine-protein kinase Kit) (c-kit) (CD117 antigen).
 GN KIT.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Fetal brain, and Term placenta;
 RX MEDLINE=8811521; PubMed=2448137;
 RA Yarden Y., Kuang W.-J., Yang-Feng T., Coussens L., Munemitsu S.,
 RA Dull T.J., Chen E., Schlessinger J., Francke U., Ullrich A.;
 RT "Human proto-oncogene c-kit: a new cell surface receptor tyrosine
 RT kinase for an unidentified ligand.";
 RL EMBO J. 6:3341-3351(1987).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=93064697; PubMed=1279499;
 RX Giebel L.B., Strunk K.M., Holmes S.A., Spritz R.A.;
 RT "Organization and nucleotide sequence of the human KIT (mast/stem
 RT cell growth factor receptor) proto-oncogene.";
 RL Oncogene 7:2207-2217(1992).
 RN [3]
 RP SEQUENCE FROM N.A.

RX MEDLINE=97179223; PubMed=9027509;
RA Andre C., Hampe A., Lachaux P., Martin E., Wang X.P., Manus V.,
RA Hu W.X., Galibert F.;
RT "Sequence analysis of two genomic regions containing the KIT and the
RT FMS receptor tyrosine kinase genes.";
RL Genomics 39:216-226(1997).
RN [4]
RP VARIANT LYS-583.
RX MEDLINE=92291284; PubMed=1376329;
RA Fleischman R.A.;
RT "Human piebald trait resulting from a dominant negative mutant allele
RT of the c-kit membrane receptor gene.";
RL J. Clin. Invest. 89:1713-1717(1992).
RN [5]
RP VARIANT LEU-584.
RX MEDLINE=92133600; PubMed=1370874;
RA Spritz R.A., Giebel L.B., Holmes S.A.;
RT "Dominant negative and loss of function mutations of the c-kit
RT (mast/stem cell growth factor receptor) proto-oncogene in human
RT piebaldism.";
RL Am. J. Hum. Genet. 50:261-269(1992).
RN [6]
RP VARIANT ARG-664.
RX MEDLINE=92020918; PubMed=1717985;
RA Giebel L.B., Spritz R.A.;
RT "Mutation of the KIT (mast/stem cell growth factor receptor)
RT proto-oncogene in human piebaldism.";
RL Proc. Natl. Acad. Sci. U.S.A. 88:8696-8699(1991).
RN [7]
RP VARIANT VAL-816.
RX MEDLINE=94013473; PubMed=7691885;
RA Furitsu T., Tsujimura T., Tono T., Ikeda H., Kitayama H.,
RA Koshimizu U., Sugahara H., Butterfield J.H., Ashman L.K.,
RA Kanayama Y., Matsuzawa Y., Kitamura Y., Kanakura Y.;
RT "Identification of mutations in the coding sequence of the proto-
RT oncogene c-kit in a human mast cell leukemia cell line causing
RT ligand-independent activation of c-kit product.";
RL J. Clin. Invest. 92:1736-1744(1993).
RN [8]
RP VARIANTS PIEBALDISM GLY-791 AND VAL-812.
RX MEDLINE=93322624; PubMed=7687267;
RA Spritz R.A., Holmes S.A., Itin P., Kuester W.;
RT "Novel mutations of the KIT (mast/stem cell growth factor receptor)
RT proto-oncogene in human piebaldism.";
RL J. Invest. Dermatol. 101:22-25(1993).
RN [9]
RP VARIANT PIEBALDISM 893-GLU-PRO-896 DEL.
RX MEDLINE=96287384; PubMed=8680409;
RA Riva P., Milani N., Gandolfi P., Larizza L.;
RT "A 12-bp deletion (781del12) in the c-kit protooncogene in a large
RT Italian kindred with piebaldism.";
RL Hum. Mutat. 6:343-345(1995).
RN [10]
RP VARIANT GIST VAL-559 DEL.
RX MEDLINE=98361155; PubMed=9697690;
RA Nishida T., Hirota S., Taniguchi M., Hashimoto K., Isozaki K.,
RA Nakamura H., Kanakura Y., Tanaka T., Takabayashi A., Matsuda H.,
RA Kitamura Y.;
RT "Familial gastrointestinal stromal tumours with germline mutation of
RT the KIT gene.";
RL Nat. Genet. 19:323-324(1998).
CC -1- FUNCTION: THIS IS THE RECEPTOR FOR STEM CELL FACTOR (MAST CELL
CC GROWTH FACTOR). IT HAS A TYROSINE-PROTEIN KINASE ACTIVITY.
CC BINDING OF THE LIGANDS LEADS TO THE AUTOPHOSPHORYLATION KIT AND
CC ITS ASSOCIATION WITH SUBSTRATES SUCH AS PHOSPHATIDYLINOSITOL
CC 3-KINASE (PI3K).
CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC tyrosine phosphate.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- DISEASE: DEFECTS IN KIT ARE THE CAUSE OF A DEVELOPMENTAL
CC ABNORMALITY KNOWN AS PIEBALDISM. AN AUTOSOMAL DOMINANT GENETIC
CC DISORDER OF PIGMENTATION CHARACTERIZED BY CONGENITAL PATCHES OF
CC WHITE SKIN AND HAIR THAT LACK MELANOCYTES.

CC -1- DISEASE: DEFECTS IN KIT ARE THE CAUSE OF A GASTROINTESTINAL
CC STROMAL TUMOR (GIST).
CC -1- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-
CC PROTEIN KINASES.
CC -1- SIMILARITY: Contains 5 immunoglobulin-like C2-type domains.
CC -1- DATABASE: NAME=NCBI; NOTE=CD guide CD117 entry;
CC WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd117.htm".
CC -1- DATABASE: NAME=Atlas Genet. Cytogenet. Oncol. Haematol.;
CC WWW="http://www.infobiogen.fr/services/chronocancer/Genes/KITID127.html".
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X06182; CAA29548.1; --
CC EMBL; X69301; CAA49159.1; JOINED.
CC EMBL; X69302; CAA49159.1; JOINED.
CC EMBL; X69303; CAA49159.1; JOINED.
CC EMBL; X69304; CAA49159.1; JOINED.
CC EMBL; X69305; CAA49159.1; JOINED.
CC EMBL; X69306; CAA49159.1; JOINED.
CC EMBL; X69307; CAA49159.1; JOINED.
CC EMBL; X69308; CAA49159.1; JOINED.
CC EMBL; X69309; CAA49159.1; JOINED.
CC EMBL; X69310; CAA49159.1; JOINED.
CC EMBL; X69311; CAA49159.1; JOINED.
CC EMBL; X69312; CAA49159.1; JOINED.
CC EMBL; X69313; CAA49159.1; JOINED.
CC EMBL; X69314; CAA49159.1; JOINED.
CC EMBL; X69315; CAA49159.1; JOINED.
CC EMBL; X69316; CAA49159.1; JOINED.
CC EMBL; U63834; AAC50968.1; --
CC PIR; S01426; THUKT.
CC HSP; P11362; IFGK.
CC Genew; HGNC:6342; KIT.
CC MIM; 164920; --
CC MIM; 172800; --
CC MIM; 606764; --
CC GO; GO:0004716; F:receptor signaling protein tyrosine kinase . . . ; TAS.
CC GO; GO:0007165; P:signal transduction; TAS.
CC InterPro; IPR007110; Ig-like.
CC InterPro; IPR003598; Ig_c2.
CC InterPro; IPR000719; Prot_kinase.
CC InterPro; IPR001824; RTKinaseIII.
CC InterPro; IPR001245; Tyr_kinase.
CC Pfam; PF00047; Ig; 1.
CC Pfam; PF00069; pkinase; 1.
CC ProDom; PD000001; Prot_kinase; 2.
CC SMART; SM00408; IGC2; 1.
CC SMART; SM00319; TykC; 1.
CC PROSITE; PS00835; IG_LIKE; 1.
CC PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
CC PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
CC PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
CC PROSITE; PS00240; RECEPTOR_TYR_KIN_III; 1.
CC Proto-oncogene; Tyrosine-protein kinase; Receptor; Transmembrane;
CC Transferase; Glycoprotein; Phosphorylation; ATP-binding; Signal;
CC Immunoglobulin domain; Disease mutation.
CC SIGNAL 1 22
CC CHAIN 23 976 MAST/STEM CELL GROWTH FACTOR RECEPTOR.
CC DOMAIN 23 520 EXTRACELLULAR (POTENTIAL).
CC TRANSMEM 521 543 POTENTIAL.
CC DOMAIN 544 976 CYTOPLASMIC (POTENTIAL).
CC DOMAIN 549 937 PROTEIN KINASE.
CC NP BIND 595 603 ATP (BY SIMILARITY).
CC BINDING 623 623 ATP (BY SIMILARITY).
CC ACT_SITE 792 792 BY SIMILARITY.
CC MOD_RES 823 823 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).

```
FT CARBOHYD 130 130 N-LINKED (GLNAC. . .) (POTENTIAL).
FT CARBOHYD 145 145 N-LINKED (GLNAC. . .) (POTENTIAL).
FT CARBOHYD 283 283 N-LINKED (GLNAC. . .) (POTENTIAL).
FT CARBOHYD 293 293 N-LINKED (GLNAC. . .) (POTENTIAL).
FT CARBOHYD 300 300 N-LINKED (GLNAC. . .) (POTENTIAL).
FT CARBOHYD 320 320 N-LINKED (GLNAC. . .) (POTENTIAL).
FT CARBOHYD 352 352 N-LINKED (GLNAC. . .) (POTENTIAL).
FT CARBOHYD 367 367 N-LINKED (GLNAC. . .) (POTENTIAL).
FT CARBOHYD 463 463 N-LINKED (GLNAC. . .) (POTENTIAL).
FT CARBOHYD 486 486 N-LINKED (GLNAC. . .) (POTENTIAL).
FT VARIANT 559 559 MISSING (IN GIST).
FT VARIANT 583 583 /FTID=VAR 007965.
FT VARIANT 584 584 E -> K (IN PIEBALDISM).
FT VARIANT 664 664 F -> L (IN PIEBALDISM).
FT VARIANT 791 791 /FTID=VAR 004105.
FT VARIANT 812 812 G -> R (IN PIEBALDISM).
FT VARIANT 816 816 R -> G (IN PIEBALDISM).
FT VARIANT 816 816 /FTID=VAR 004107.
FT VARIANT 816 816 G -> V (IN PIEBALDISM).
FT VARIANT 816 816 /FTID=VAR 004108.
FT VARIANT 816 816 D -> V (IN MAST CELL LEUKEMIA;
FT VARIANT 816 816 CONSTITUTIVELY ACTIVATED).

Query Match 23.7%; Score 1251; DB 1; Length 976;
Best Local Similarity 32.7%; Pred. No. 4.7e-78;
Matches 324; Conservative 166; Mismatches 352; Indels 148; Gaps 34;

QY 47 VCKSSYPWSSPDLGALRPOSSGTVYEAADVVSASITLOVLVDAPGNISCLWV 106
DB 20 VOTGSSQSVSG-BPSPPSIHPGKSDL-----VRVDEIRLLCTDGFVK-WT 67
QY 107 FKHSLLNCPHPDLQNRGVVSVILKWTGTQAGEYLLFIQSEATN---YTILFTWSIRNT 163
DB 68 FE-----ILDTENKQNEWT-EKAEATNTGKYCTNKHGLSNS 106
QY 164 L-----LYTLRPPYPRKMNQDALVCISEVPPIVEWVLCDSQGESKBE----- 209
DB 107 IYVVRDPKLFVDRSLYK-EDNDTLVRCPLTDP-EVTNYSKGCOCGKPLPKDLRFP 164
QY 210 ---SPAUVKKEKVLHFGTGTDIRCCARNELGRECTRLFTIDL-----NOTPQTLPQ--L 260
DB 165 DPKAGIMIKSVKRAVHRLC---LHCSVDQEGKSVLSEKFIKVRPAKFAVVPVSVKASY 221
QY 261 FLKVGEPMLIRKAVHNVHFGLTWELN---KALEEGNYFEMSTYSTNTRTMIRILFAFV 317
DB 222 LLRGEFTVCTIKDVSSVSVTWKRENSQYKQKYNHWHGDFNYERQAT-----LTI 277
QY 318 SSVARNDTGYTTCSSSKHPSQALVT---IVGKGFNA-----TNSSEDEYIDQYE 365
DB 278 SSARVNDSGVFCYANNTPGSANVTTLVVDKGFNFIMPINTTVFVNDGENVDL--- 333
QY 366 EPCFSVRKAYPQ-IRCTWTFERSKP--PCQKGLDNGYSISKFNKH-----OPGE 415
DB 334 ----IVEYFAFPKPEHQIYNNRTFTDKWEDPKSESNIRYVSELHLTRLKGTGEGT 389
QY 416 YIFHAENDDAQTKFTMLIRKPOVLA--EASASQACFSQDGYPLPSPWTKKCDKSPN 473
DB 390 YTFVLSNDSVNAIAFNVTYNTKPELITVDRLVNGMLQCVAAGFPPEPTIDWFCPTEQR 449
QY 474 CTEETE-GVMNRKANRKYFGQWVSSTLNMGEBAIKGFLVKCCAYNSLGTSCETILLNSP 532
DB 450 CSASVLPVDVQTLNSSGPPFKLVQSSIDSSAFKNGTVECKAYNDVGKT--SAYFN-- 505
QY 533 GPFPIQDN-----ISFYAIGVCLLFIVVLTLICHYKQKFRYESQLQM 578
DB 506 --FAFKGNKKEQIHPHTLFTPLLIIGFVIVAGMCIIIVMILT-----YKYLQKPMYEQWV 559
QY 579 V-QVTGSSDNEYFYVDREYEDLWKEPPRENLERGKVLGSGAFGKVMNATYAGISKTV 637
DB 560 VEEING---NNYVYIDPTQLPDYDHRKWEPPRNRLSFGKTLGAGAFGKVEATAYGLIKSDA 616
```

```
QY 638 STQAVKMLKEKADSSEREALMSLKMMTOLGSHENIVNLLGACTLSGPIYLIFEYCCYG 697
DB 617 AMTAVKMLKPSAHLTEREALMSLKVLKSYLGNHWNIVNLLGACTIGGTLVITECCYG 676
QY 698 DLLNLYLRKREKPHRTWTE-----IFKEHNFSPYTFQSHPNSSM---PGSREVOIHPD 748
DB 677 DLLNLFRRKRDSFICSKQEDHAAALYKNLLHSKSSCSDSSTNEYMDMKPGVSYV-VPTK 735
QY 749 SQOISGLHNSPHSDETEYENQKLEEBEDLNVLTFEDLLCFAYOVAKMELEPKSCV 808
DB 736 ADKRSVRIGSY-----IERDVTPTAIMEDDEL-ALDLELLSFSYQVAKMAFLSKNCI 789
QY 809 HDLAARNVLVTHGKVKVCKIDFGLARDIMSDSNVYVGRNARLPVKWMAPESLPEGIVYTIK 868
DB 790 HDLAARNILLTHGRITKICDFGLARDIKDSDNVYVGRNARLPVKWMAPESFNCVYTFE 849
QY 869 SDVWSYGILLWEIFSLGVNYPYGPVDFANFYKLQNGFQDQOPFYATEBIYIIMQSCWAF 928
DB 850 SDVWSYGIFLWELFSLGSSPYGMPVDSKFYKWKIEGPRMLSPHEHAPAEWYDIMKTCWDA 909
QY 929 DSRKPSPNLTSLFLGCLQADAEEMVQNV 958
DB 910 DPLKPTFKQIVOLIEKOISESTNHIYSNL 939

RESULT 7
KIT_CANFA
ID KIT_CANFA STANDARD; PRT; 975 AA.
AC O97799;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Mast/stem cell growth factor receptor precursor (EC 2.7.1.112) (SCFR)
DE (Proto-oncogene tyrosine-protein kinase Kit) (c-kit).
GN KIT.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99142897; PubMed=9989791;
RA Ma Y., Longley B.J., Wang X., Blount J.L., Langley K., Caughey G.H.;
RT "Clustering of activating mutations in c-KIT's juxtamembrane coding
RT region in canine mast cell neoplasms."
RL J. Invest. Dermatol. 112:165-170(1999).
CC -!- FUNCTION: THIS IS THE RECEPTOR FOR STEM CELL FACTOR (MAST CELL
CC BINDING FACTOR). IT HAS A TYROSINE-PROTEIN KINASE ACTIVITY.
CC BINDING OF THE LIGANDS LEADS TO THE AUTOPHOSPHORYLATION KIT AND
CC ITS ASSOCIATION WITH SUBSTRATES SUCH AS PHOSPHATIDYLINOSITOL
CC 3-KINASE (PI3K).
CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC tyrosine phosphate.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-
CC PROTEIN KINASES.
CC -!- SIMILARITY: Contains 5 immunoglobulin-like C2-type domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF044249; AAD02327.1; -.
CC HSSP; P11362; 1FGK.
CC InterPro; IPR007110; Ig-like.
CC InterPro; IPR003599; Ig.
CC InterPro; IPR003006; Ig_MHC.
CC InterPro; IPR000719; Prot_Kinase.
CC InterPro; IPR001824; RTKinaseII.
```

DR	InterPro; IPR001245; Tyr_kinase.	QY	471	SPNTEEI-TEGVNRKANRKFVGVQWSSSTLNMSEAIKGLFKVCCAYNSLGTSCETILL	529
DR	Pfam; PF00047; IG; 2.	Db	450	EORCSVPVIGPMDVQVQNSLSLSPSGKLVVQSSIDISAFKHNGTVECRATNVGRS--SAFF	507
DR	ProDom; PD000001; Prot_kinase; 1.	QY	530	NSPGPPFPFIQD-----NISFYATIGVCLLFIVVLTLLICHYKQKOFYESQLOQV	579
DR	SMART; SM00409; IG; 3.	Db	508	N----FAFKEQIHPHTLFTPLLIGFVIAAGMCIIVMLT-----YKYLQKPMYEQWVV	559
DR	SMART; SM00219; TyrK; 1.	QY	580	-QVTGSSDNEYVDVDFREYEDLKWEPPREMLFQKVLGSGAFGKMMATAYGISTGVS	638
DR	PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.	Db	560	EBING---NNVYIDPTQLPYDHWKWEPPRNLRSFGKTLGAGAFGVBEATAYGLKSDAA	616
DR	PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.	QY	639	IQVAVMLKEKADSSERELMSKMTQLGSHENIVNLGACTLSGPIYLIFEFCCYGD	698
DR	PROSITE; PS00240; RECEPTOR_TYR_KIN_III; 1.	Db	617	MTAVVLMKPSAHLTEREALMSKLVLSYLGNNHNVNLLGACTVGGPTLVITECCYGD	676
KW	Proto-oncogene; Tyrosine-protein kinase; Receptor; Transmembrane;	QY	699	LLNLYRSKREKPFHRTWTE-----IFKEHNFSEFYTFQSHPNSSM---PGSREVOIHPDS	749
KW	transferase; Glycoprotein; Phosphorylation; ATP-binding; Signal;	Db	677	LLNFLRRRDSFICSKQEDHGEVALYKLLHSSKSSCSDSSTNEYMDKGVSVV-VPTKA	735
FT	Immunoglobulin domain.	QY	750	DQISGLHGNFSHSEDEIYENQKLEEDLNVLFFDLFCFAYQVAKMEFLFKSCVH	809
FT	SIGNAL 1 24	Db	736	DKRSARISY-----IERDVTVAIMEDDEL-ALDLEDLSFSYQVAKGMFLASKNCIH	789
FT	CHAIN 25 975	QY	810	RDLAARNVLVTHGKVKIKCDFGLARDIMSDSNVYVRGHAFLPVKMAPESLFEGITIKS	869
FT	DOMAIN 25 519	Db	790	RDLAARNLLTHGRITKICDFGLARDIKNDNSVYVVGARLPVKWMAPEISFNCVYTFES	849
FT	TRANSMEM 520 542	QY	870	DWMSGIILLWEISGLVGNVPYGPVDFANFYKLIQNGKMDQDPYATEEYIIMOSCAFD	929
FT	DOMAIN 543 975	Db	850	DWMSGIFELWEISGLVGNVPYGPVDFANFYKLIQNGKMDQDPYATEEYIIMOSCAFD	909
FT	NP_BIND 594 602	QY	930	SRKPSFENLTFSLGCLADAEAEAMYNQV 958	
FT	NP_BIND 622 622	Db	910	PLKRTPEKQIVQLIEKQISDSSTNHIYSNL 938	
FT	ACT_SITE 791 791	QY			
FT	MOD_RES 822 822	Db			
FT	CARBOHYD 96 96	QY			
FT	CARBOHYD 132 132	Db			
FT	CARBOHYD 147 147	QY			
FT	CARBOHYD 286 286	Db			
FT	CARBOHYD 296 296	QY			
FT	CARBOHYD 303 303	Db			
FT	CARBOHYD 355 355	QY			
FT	CARBOHYD 370 370	Db			
FT	CARBOHYD 403 403	QY			
FT	CARBOHYD 466 466	Db			
FT	CARBOHYD 489 489	QY			
FT	CARBOHYD 975 975	Db			
FT	SEQUENCE 975 AA; 109335 MW; 8F570BDBF05B1CB CRC64;	QY			
QY	Query Match 23.7%; Score 1247.5; DB 1; Length 975;				
Db	Best Local Similarity 33.7%; Pred. No. 8.2e-78;				
QY	Matches 333; Conservative 161; Mismatches 346; Indels 149; Gaps 39;				
Db	47 VKSSSPVMSPEPDLGALRPSQSGTVEAANVEDVVSASITLQVLDAGNLSCLVW 106				
QY	22 VRTGSQSPVSGPESLP-SIHPAKS-----ELIVSGDELRLCTDGFVK--WT 69				
QY	107 PKH-SSLNCPHFQDLQNRGVSMVLKMTQAGEYLLFIQSEATNYTIL-----FT 157				
Db	70 ETLQQLNETH-----NEWITE-----KAEAG-----HTGNYTCTNRDGLRSY 110				
QY	158 VSIRN-TLLYTLRRPYFRKMNQDALVCISESVPEPIVWVLCDSQGESKEESP 213				
Db	111 VFVRDPKFLVDPLPYGK-EGNDTLVRCPLTDPE-VTVNLSRGCEGKPLPKDLTFVADP 168				
QY	214 -----VKKEKVLHFLGTDIRC--CARNELGRE-CTRLFTTD-----LNQTFQTLPOI- 260				
Db	169 KAGITIRNVKREYHRL-----CLHGSADQKGRVTLKSKFTLVRAAIRAVPVSVSKTS 222				
QY	261 -FLKVGPELWIRCKAVHNVHGLFTWLEN-----KALESGNYFEMSTYSTNRTMIRLFA 315				
Db	223 SLKGEAFSVNCFIKDVSSFVDSMMWIKENSQOQTNAQTSNMHHGDFNFEQEKLI----- 279				
QY	316 FVSSVARNDTGVTYSSSKHPSOSALVT---IVGKGFINA-----TNSSEYEDIQ 363				
Db	280 -ISSARVNDGVMFCYANNVTFGSANVTTLLEVVDKGFINIFPMWSTTFIVNDQNVLD-- 336				
QY	364 YEEFCFSVRFKAYPQ-IRCTWTFSRKSPCEQKGL---DNGYSISKFNKH-----Q 412				
Db	337 -----IVEYAYPKHQOQWYNNRTFTDKWEDYKSDNESNI-RVYSELHLTLKNGNE 389				
QY	413 PEYIFHAENDDAQFTKMTLNIARRPQVLAESASQA--SCFSDGYPLPSWTWKCSDK 470				
Db	390 GGTYYTFQVNSDVSNSVTFNVYNTKPEILTHESLNGMLQCVVAGFPFPAVGVYFCPGA 449				

RESULT 8
KFMS FELCA
ID KFMS FELCA STANDARD; PRT; 980 AA.
AC P13369;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Macrophage colony stimulating factor 1 receptor precursor (CSF-1-R) (EC 2.7.1.112) (fms proto-oncogene) (c-fms).
GN CSF1R OR FMS.
OS *Felis silvestris catus* (Cat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
OX NCBI TaxID=9685;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89077553; PubMed=2849512;
RA Woolford J., McAuliffe A., Kohrschneider L.R.;
RT "Activation of the feline c-fms proto-oncogene: multiple alterations are required to generate a fully transformed phenotype.";
RL Cell 55:965-977(1988).
CC -!- FUNCTION: THIS PROTEIN IS THE RECEPTOR FOR CSF-1, IT IS A PROTEIN
CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC -!- TYROSINE-KINASE TRANSMEMBRANE RECEPTOR.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-
CC -!- PROTEIN KINASES.
CC -!- SIMILARITY: Contains 5 immunoglobulin-like C2-type domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

tyrosine-substrate interaction site.";
 Oncogene 10:369-379(1995).
 -!- FUNCTION: THIS IS THE RECEPTOR FOR STEM CELL FACTOR (MAST CELL GROWTH FACTOR). IT HAS A TYROSINE-PROTEIN KINASE ACTIVITY.
 BINDING OF THE LIGANDS LEADS TO THE AUTOPHOSPHORYLATION KIT AND ITS ASSOCIATION WITH SUBSTRATES SUCH AS PHOSPHATIDYLINOSITOL 3-KINASE (PI3K).
 -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein tyrosine phosphate.
 -!- SUBCELLULAR LOCATION: Type I membrane protein.
 -!- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-PROTEIN KINASES.
 -!- SIMILARITY: Contains 5 immunoglobulin-like C2-type domains.
 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).
 EMBL; S76596; AAB33207.1; --
 HSP; P11362; 1FGK
 InterPro; IPR007110; Ig-like.
 InterPro; IPR003599; Ig.
 InterPro; IPR003006; Ig_MHC.
 InterPro; IPR000719; Prot_kinase.
 InterPro; IPR01824; RTKinaseIII.
 InterPro; IPR001245; Tyr_pkinase.
 Pfam; PF00047; Ig; 1.
 Pfam; PF00069; pkinase; 1.
 ProDom; PD000001; Prot_kinase; 2.
 SMART; SM00409; IG; 3.
 SMART; SM00219; Tyrc; 1.
 PROSITE; PS50835; IG_LIKE; 2.
 PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
 PROSITE; PS00240; RECEPTOR_TYR_KIN_III; 1.
 K0 Proto-oncogene; Tyrosine-protein kinase; Receptor; Transmembrane;
 KW Transferase; Glycoprotein; Phosphorylation; ATP-binding; Signal;
 KW Immunoglobulin domain.
 SIGNAL 1 22 POTENTIAL.
 CHAIN 23 978 MAST/STEM CELL GROWTH FACTOR RECEPTOR.
 DOMAIN 23 521 EXTRACELLULAR (POTENTIAL).
 TRANSMEM 522 544 POTENTIAL.
 DOMAIN 545 978 CYTOPLASMIC (POTENTIAL).
 DOMAIN 590 938 PROTEIN KINASE.
 NP BIND 596 604 ATP (BY SIMILARITY).
 BINDING 624 624 ATP (BY SIMILARITY).
 ACT SITE 793 793 BY SIMILARITY.
 MOD_RES 824 824 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
 CARBOHYD 94 94 N-LINKED (GLCNAC. .) (POTENTIAL).
 CARBOHYD 145 145 N-LINKED (GLCNAC. .) (POTENTIAL).
 CARBOHYD 284 284 N-LINKED (GLCNAC. .) (POTENTIAL).
 CARBOHYD 294 294 N-LINKED (GLCNAC. .) (POTENTIAL).
 CARBOHYD 301 301 N-LINKED (GLCNAC. .) (POTENTIAL).
 CARBOHYD 321 321 N-LINKED (GLCNAC. .) (POTENTIAL).
 CARBOHYD 353 353 N-LINKED (GLCNAC. .) (POTENTIAL).
 CARBOHYD 368 368 N-LINKED (GLCNAC. .) (POTENTIAL).
 CARBOHYD 401 401 N-LINKED (GLCNAC. .) (POTENTIAL).
 CARBOHYD 464 464 N-LINKED (GLCNAC. .) (POTENTIAL).
 CARBOHYD 487 487 N-LINKED (GLCNAC. .) (POTENTIAL).
 SEQUENCE 978 AA; 109449 MW; 6D45472E07440E6B CRC64;

Query Match 23.2%; Score 1224.5; DB 1; Length 978;
 Best Local Similarity 32.9%; Pred. No. 3.1e-76;
 Matches 325; Conservative 167; Mismatches 352; Indels 145; Gaps 35;
 47 VCKSSSYPMVSSPEDLGC-ALRPQSGTYVAAAEVDVVSASITQLVLVDAPGNISCLW 105
 20 YOTGSSQP--SASPGWMSLPSIHPATS-----ELIVSAGDEIRLLCTDGFVK--W 66

106 VPKHSSLNCQHFQDLQNGVGVSWILKMTQAGCYLLFIOSEAT---NYTIL----- 155
 67 TFE-----TLGQSSEITHNEWIT-EKAEATNTGNTTNGGGLSS 105
 156 -FTYSIRN-TLLYTLRRPYFRKMNQDALVICISESVPEPIVVEWVLCDSOGESKEESPAV 213
 106 SIYVVRDPAPKLPLVDLPYLGK-EDHDTLVRCPLTDPB-VTNYSLRGCGKPLPLDLTFV 163
 214 -----VKKEKVLHFLGTDIRCCA-----RNLGRECTRLFTIDLNTQTPTTLPLQL- 260
 164 TDPKAGITIRNVKREYHRLC---LHCSADRGKSVLSKFTLKRAAIRAAPVSVSKAS 220
 261 -FLKVGEPILTRCKAVHVNHGFLTWELNKALESGNYFEMFSTYSTRNTRMIRILFAFVS 319
 221 HLLRGEFEFVNCILTKDVSSVSDSMWIKENSPQNAQPSNSWHQDQNFVFOERLITSS 280
 320 VARNDTGYVTCCSSKHPQSALVT---IVGKGFINA-----TNSSEYDEIDQYESEF 367
 281 ARVNDGSGVMFCVANTFSGANVTTLVAVAKGFNIFPMWNTTIFVNDGENDVL----- 334
 368 CFSVRFKAYPQ-----IRCTWTFSRKSPCEQKGLDNGYSISKFCNHKH-----Q 412
 335 --IVEYEAVPKPEHORWVVMNRTLTDKWEDPKS---DNESNI-RYVSELHLTRLKGNE 387
 413 PGEYIFHAENDDAQFTKMTFLNIRKPKQVLAAS--ASQASCFSDGYPPLSPSTWTKCSDK 470
 388 GGTYYTFQVNSDVNSVTLNVVYVTKPILTHESLVGILQCLVAGFPPEPTVDWVFCPGA 447
 471 SPNCTEEITE-GVMNRKANRKVFGOWSSSTLNLMSEAIGPLVCKCAYNSLGTSCBTILL 529
 448 EORCPVPVGPLDVQMNSVSPSGKLVQVSSIDYSAFKINGTVECRASNNGVT--SAFF 505
 530 NSPGPFPIQDN-----ISFYATIGVCLLFIIVVLTLLIHKYKQKPYESQ 575
 506 N-----FAFKGSKSQMHPTLFTPLLIGPVIAAGMCIIVMLT---YKYLQKPMVEVQ 557
 576 LQMV-QVTGSSDNEYFYDFREYEDLKWEPRENLEFGKVLGSGAFGKMMATAYGISK 634
 558 WKVVEEING---NNYVYIDPTOLPYDHKWEFPNRLSFGKTLGAGAFGKVEATAYGLIK 614
 635 TGVSIOQAVKMLKEKADSSREALMSLKMOTQLGSHENIVNLGACTLSGPTLYLIFYEC 694
 615 SDAAVTAVKMLKPSAHUTEREALESKLVSLYGNHMINVNLGACTVGGPTLVITEYC 674
 695 CYGDLNLYLRKREKFRHTWTEIFKEHNFSPYFT-FQSHPSNSMPGSRV-VQIHPDSQI 752
 675 CYGDLNLYLRKREKFRHTWTEIFKEHNFSPYFT-FQSHPSNSMPGSRV-VQIHPDSQI 732
 753 SGLHNSFHSF---DEIEYENOKRLEEBEDLVNLTFFEDLLCFAYQVAKGMEFLFKSCVH 809
 733 VPTKADKRSARIGSYIERDVTTPAIMEDEL-ALDLEDLLSFSYQVAKGMFLASKNCIH 791
 810 RLAAARNVLVTHGKVKIKCDFGLARDIMSDSVVVGARLNVKMAPESLPEGIVTIS 869
 792 RLAAARNILLTHGRITIKCDFGLARDIKNDNSVYVKGARLNVKMAPESINFCVTFES 851
 870 DVWSYGILLWEIFSLGVNPPYGPVVDANFYKLIQNGFKMDQPPYATEEYIIMQSCWAF 929
 852 DVWSYGIFLWELFSLGSSPYPCMPVDSKEYKMKEGFRMLSEHAPAEWYDINKTCWAD 911
 930 SRKRPSFNLTSFLGCLQADAEAEANYQNV 958
 912 PLKRPTFKQVQLIEKQISDSTNHIYSNL 940

RESULT 10
 KFWMS FSVM
 ID KFWMS FSVM STANDARD; PRT; 978 AA.
 AC P00545; Q86597;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-NOV-1991 (Rel. 20, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)

DE Tyrosine-protein kinase transforming protein fms (BC 2.7.1.112).

GN V-FMS

OS Feline sarcoma virus (strain McDonough).

OC Viruses; Retroviral viruses; Retroviridae; Mammalian type C retroviruses.

OC NCBI_TaxID=111778;

[1]

RP SEQUENCE FROM N.A.

RA MEDLINE=84119469; PubMed=6582485;

RA Hampe A., Gobet M., Sherr C.J., Galibert F.;

RT "Nucleotide sequence of the feline retroviral oncogene v-fms shows

RT unexpected homology with oncogenes encoding tyrosine-specific protein

RT kinases.";

RL Proc. Natl. Acad. Sci. U.S.A. 81:85-89(1984).

RN [2].

RP REVISIONS, SEQUENCE FROM N.A.

RA MEDLINE=92015516; PubMed=1833563;

RA Smola U., Hennig D., Hadwiger-Pangmeier A., Schuetz B., Pfaff E.,

RA Niemann H., Tamura T.;

RT "Reassessment of the v-fms sequence: threonine phosphorylation of the

RT COOH-terminal domain.";

RL J. Virol. 65:6181-6187(1991).

CC -!- FUNCTION: V-FMS IS DERIVED FROM THE RECEPTOR FOR COLONY

CC -!- STIMULATING FACTOR 1 (CSF-1).

CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein

CC tyrosine phosphate.

CC -!- SUBCELLULAR LOCATION: Type 1 membrane protein.

CC -!- MISCELLANEOUS: THIS PROTEIN IS SYNTHESIZED AS A GAG-FMS

CC POLYPEPTIDE.

CC -!- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-

CC PROTEIN KINASES.

CC -!- SIMILARITY: Contains 5 immunoglobulin-like C2-type domains.

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC

CC EMBL; K01643; AAA43045.1; -

CC EMBL; S59588; AAB20028.1; -

CC PIR; A00654; TMVMD.

CC

CC HSP; P11362; IFCK.

CC InterPro; IPR007110; Ig-like.

CC InterPro; IPR003598; Ig_c2.

CC InterPro; IPR003006; Ig_MHC.

CC InterPro; IPR000719; Prot_kinase.

CC InterPro; IPR001824; RTKinaseIII.

CC InterPro; IPR001245; Tyr_pkinase.

CC Pfam; PF00047; Ig_3.

CC Pfam; PF00069; pkinase; 1.

CC ProDom; PD000001; Prot_kinase; 2.

CC SMART; SM00408; IgC2; 1.

CC SMART; SM00219; Tyrc; 1.

CC PROSITE; PS00835; IG LIKE; 3.

CC PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.

CC PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.

CC PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.

CC PROSITE; PS00240; RECEPTOR_TYR_KIN_III; 1.

CC PolyProtein; Tyrosine-protein Kinase; Oncogene; Transferase; Receptor;

CC transmembrane; Glycoprotein; Phosphorylation; ATP-binding;

CC Immunoglobulin domain; Repeat.

FT DOMAIN 1 543 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 544 568 POTENTIAL.

FT DOMAIN 569 978 CYTOPLASMIC (POTENTIAL).

FT DOMAIN 55 134 IG-LIKE C2-TYPE 1.

FT DOMAIN 141 231 IG-LIKE C2-TYPE 2.

FT DOMAIN 236 331 IG-LIKE C2-TYPE 3.

FT DOMAIN 333 431 IG-LIKE C2-TYPE 4.

FT DOMAIN 434 533 IG-LIKE C2-TYPE 5.

FT DOMAIN 613 942 PROTEIN KINASE.

FT NP_BIND 619 627 ATP (BY SIMILARITY).

FT BINDING 647 647 ATP (BY SIMILARITY).

FT DISULFID 76 118 POTENTIAL.

FT DISULFID 151 211 POTENTIAL.

FT DISULFID 258 312 POTENTIAL.

FT DISULFID 451 516 POTENTIAL.

FT MOD_RES 841 841 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).

FT MOD_RES 973 973 PHOSPHORYLATION.

FT CARBOHYD 79 79 N-LINKED (GLCNAC...) (POTENTIAL).

FT CARBOHYD 107 107 N-LINKED (GLCNAC...) (POTENTIAL).

FT CARBOHYD 128 128 N-LINKED (GLCNAC...) (POTENTIAL).

FT CARBOHYD 187 187 N-LINKED (GLCNAC...) (POTENTIAL).

FT CARBOHYD 309 309 N-LINKED (GLCNAC...) (POTENTIAL).

FT CARBOHYD 320 320 N-LINKED (GLCNAC...) (POTENTIAL).

FT CARBOHYD 336 336 N-LINKED (GLCNAC...) (POTENTIAL).

FT CARBOHYD 369 369 N-LINKED (GLCNAC...) (POTENTIAL).

FT CARBOHYD 444 444 N-LINKED (GLCNAC...) (POTENTIAL).

FT CARBOHYD 511 511 N-LINKED (GLCNAC...) (POTENTIAL).

FT CARBOHYD 524 524 N-LINKED (GLCNAC...) (POTENTIAL).

FT CONFLICT 714 714 L -> P (IN REF. 1).

FT CONFLICT 971 978 QRTPPVAR -> RGPL (IN REF. 1).

SQ SEQUENCE 978 AA; 108491 MW; 4C7CAC4835185EBF CRC64;

Query Match 32.2%; Score 1224; DB 1; Length 978;

Best Local Similarity 32.6%; Pred. No. 3.4e-76;

Matches 334; Conservative 173; Mismatches 326; Indels 192; Gaps 39;

QY 65 CALRPOSSG-----TVYEAANE-----VDVASITLQVLVDAPGNISCL 104

DB 28 CCLPTEAMPRALVLLMATAWHAQGVPIQSPGPELVPEPTVTTLRCV-----GNGSVE 83

QY 105 W---VFKHSSNCQP-----HFDLQNRGVSVMLKWTQTQAGE-----YLLFIQ 146

DB 84 WDGPISPHNLDDPPSILTNATFQNTGYHC-----TEPGNPGGNATHLYVK 136

QY 147 SEATNYTL--FTVSIRNTLLYLRPFYRKMENQDALV-CISESVPEPIVE--WVLCD 200

DB 137 DPAPRWKVLAEQTVV-----LEGQDALLPCL--LTPALEAGVSLVR 176

QY 201 SQGESCKEE-----SP---AVVKEEKLHELFGDTRCCARNELGRETRL-----FTI 247

DB 177 VRGRPVLRQTVNPSFSPHGFTHKAKFIENHYV---QCSARVD-GRVTSMGLWLVKQV 231

QY 248 DLNQTPQTL-POLELVK-GEPLWIRCKAVHVGHGLTWELENKALEGN-----YFMS 301

DB 232 DISGPATLTLEPAELVRIQGEAAQIVCSASNDVNFV-----SLRHGDTKLITSQOS 284

QY 302 TYSTNRTMIRILFAFVSSVARNDTGYTCSSSK---HPSQSALVTIVGKGFNATN--SSE 357

DB 285 DFHDNRYQ-KVLTMLNDHVSFQDAGNYSCATNANGNHSASWFRVVEGAYNLTBEQSL 343

QY 358 DYEIDQYEEFCFSVRKAYPQIRC-TWTFSRKSPFCEQKGLD-----NGYSISKFCN--- 408

DB 344 LQEVTVGEKVDLQVKVEAYPGLESFNWIV-LGPFSDYQDKLDFVTKIDRYTSTLSLPR 402

QY 409 -HKQPGYIYIHAENDDAQTFQMTINTRRKQVLAESAASOAS-----CFSDGYPIPSWT 463

DB 403 LKRSSEGRYSFLARNAGQGNALTPELTLYRPEVVRTWTLINGSQDITLLCEASGYPPSVT 462

QY 464 WKCKSDKSPNCTEE---ITEGVWNRKANRKYGVQWVSSSTLNWSEAIKFLVKCCAYNSL 520

DB 463 WVQCKSHYDRCDSEAGLVLESHSVLSQVPEYEVIVHSLAIGTLNHRNTECFRAFNSV 532

QY 521 GTSCEITLLNSPGPPFFQDNISFYATIGVC---LLFIVLTLTLLCHYKQKQFVRESQIQ 577

DB 523 GNSQTFWPIGIAHTPLDELLFTFVLLTCSINALLLLLLLLLLLYKKYKQKQVQRWK 582

QY 578 MVQVTGSSDNRYFYVDREYDYDLKWEFPRENLEFGKVLGSGAFGKVMNATYGIKTVG 637

DB 583 IIE--SYEGNSYTFIDPTQLPYNEKWEFPERNLQFGKTLGTGAFGKVEATEAFGLKEDA 640

QY 638 SIQVAVKMLKEKADSSEREALMSELKMTQLGSHENIVNLLGACTLSGIYLIFFECCYCG 697

DB 641 VLKVAVRMLKTAHADEKELMSELKIMSHLQGHENIVNLLGACTHGGPVLVITECCYCG 700

QY 698 DLLNVLRSKRE-----KFHRTWTETFEKHNPSFYPTTQSHPSNM 737
 DB 701 DLLNFRQAEAMLSVGDPEAGYKNIHLEKKYVRDSDGFS----- 747
 QY 738 PGSREVOIHPDSQISGLHGNFSHSEDELEYENQKLEEEEDLNVLTPEDLLCFAYOVAK 797
 DB 748 --SQGVDTIVMRPVSSTSSNDSFSEEDLGKEDGRPLE-----LRDLLHFSQVAQ 796
 QY 798 GMEFLFKSCVHRDLAARNVLTHGKVKICDFGLARDIMSDSNVYVGRNARLPVKWMA 857
 DB 797 GMFLASKNCIHRDVAARNVLTSGRVAKIGDFGLARDIMSDSNVYVGRNARLPVKWMA 856
 QY 858 ESLPFGIYTIKSDVWSYGILLWEIFSLGNVPIGIPVDANFYKLTQNGFMDQPFYATBE 917
 DB 857 ESIFCVYTVQSDVWSYGILLWEIFSLGNVPIGIPVDANFYKLTQNGFMDQPFYATBE 916
 QY 918 IYIMQSCWAFDSRKRPSFPNLTSLFCGLADAEARMYQNVQGRV---SECPTTYQNR-- 972
 DB 917 IYIMQSCWAFDSRKRPSFPNLTSLFCGLADAEARMYQNVQGRV---SECPTTYQNR-- 972
 QY 973 RPFSSR 977
 DB 968 RFWQR 972

RESULT 11
 ID_KFMS HUMAN STANDARD; PRT; 972 AA.
 AC P07333,
 DT 01-APR-1988 (Rel. 07, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Macrophage colony stimulating factor I receptor precursor (CSF-1-R)
 DE (EC 2.7.1.112) (fms proto-oncogene) (c-fms) (CD115 antigen).
 GN CSF1R OR FMS.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89239490; PubMed=2524025;
 RA Hampe A., Shamoon B.M., Gobet M., Sherr C.J., Galibert F.;
 RT "Nucleotide sequence and structural organization of the human FMS
 RT proto-oncogene";
 RL Oncogene Res. 4:9-17(1989).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=86175013; PubMed=2421165;
 RA Coussens L., van Beveren C., Smith D., Chen E., Mitchell R.L.,
 RA Isacke C.M., Verma I.M., Ullrich A.;
 RT "Structural alteration of viral homologue of receptor proto-oncogene
 RT fms at carboxyl terminus";
 RL Nature 320:277-280(1986).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Placenta;
 RX MEDLINE=97179223; PubMed=9027509;
 RA Andre C., Hampe A., Lachaux P., Martin E., Wang X.P., Manus V.,
 RA Hu W.X., Galibert F.;
 RT "Sequence analysis of two genomic regions containing the KIT and the
 RT FMS receptor tyrosine kinase genes";
 RL Genomics 39:216-226(1997).
 RN [4]
 RP SEQUENCE OF 1-16 FROM N.A.
 RX MEDLINE=89261741; PubMed=2524648;
 RA Visvader J., Verma I.M.;
 RT "Differential transcription of exon 1 of the human c-fms gene in
 RT placental trophoblasts and monocytes";
 RL Mol. Cell. Biol. 9:1336-1341(1989).
 RN [5]
 RP SEQUENCE OF 1-16 FROM N.A.

RX MEDLINE=86281820; PubMed=3525854;
 RA Wheeler E.F., Rousset M.F., Hampe A., Walker M.H., Fried V.A.,
 RA Look A.T., Rettemier C.W., Sherr C.J.;
 RT "The amino-terminal domain of the v-fms oncogene product includes a
 RT functional signal peptide that directs synthesis of a transforming
 RT glycoprotein in the absence of feline leukemia virus gag sequences.";
 RL J. Virol. 59:224-233(1986).
 CC -!- FUNCTION: THIS PROTEIN IS THE RECEPTOR FOR CSF-1, IT IS A PROTEIN
 CC TYROSINE-KINASE TRANSMEMBRANE RECEPTOR.
 CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
 CC tyrosine phosphate.
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -!- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-
 CC PROTEIN KINASES.
 CC -!- SIMILARITY: Contains 5 immunoglobulin-like C2-type domains.
 CC -!- DATABASE: NAME=PRO; NOTE=PROW 1:13-21(2000);
 CC WWW="http://www.ncbi.nlm.nih.gov/prov/guide/1999017666.g.htm".
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M25786; AAA58421.1; -
 DR EMBL; U63963; AAA51696.1; -
 DR EMBL; M14002; AAA35849.1; -
 DR EMBL; X03663; CAA27300.1; -
 DR F01; S08123; TVHUND.
 DR HSP; P11362; IFGK.
 DR GENE; HGNC:2433; CSF1R.
 DR MIM; 164770; -
 DR GO; GO:0005887; C:integral to plasma membrane; TAS.
 DR GO; GO:0005011; F:macrophage colony stimulating factor recept. . . ; TAS.
 DR GO; GO:0006960; P:antimicrobial humoral response (sensu inver. . . ; TAS.
 DR GO; GO:0008283; P:cell proliferation; TAS.
 DR GO; GO:0007275; P:development; TAS.
 DR GO; GO:0007165; P:signal transduction; TAS.
 DR InterPro; IPR007110; Ig-like.
 DR InterPro; IPR003598; Ig_C2.
 DR InterPro; IPR003006; Ig_MHC.
 DR InterPro; IPR000719; Prot_kinase.
 DR InterPro; IPR001824; RTKinaseIII.
 DR InterPro; IPR001245; Tyr_kinase.
 DR Pfam; PF00047; ig; 4.
 DR Pfam; PF00069; pkinase; 1.
 DR ProDom; PD000001; Prot_kinase; 2.
 DR SMART; SM00408; IGC2; 1.
 DR SMART; SM00219; Tyrc; 1.
 DR PROSITE; PS0835; IG LIKE; 3.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
 DR PROSITE; PS00240; RECEPTOR_TYR_KIN_III; 1.
 KW Proto-oncogene; Tyrosine-protein kinase; Receptor; Transmembrane;
 KW Transferrase; Glycoprotein; Phosphorylation; ATP-binding; Signal;
 KW Immunoglobulin domain; Repeat; Polymorphism.
 FT SIGNAL 1 19
 FT CHAIN 20 972
 FT MACROPHAGE COLONY STIMULATING FACTOR I
 FT RECEPTOR.
 FT EXTRACELLULAR (POTENTIAL).
 FT POTENTIAL.
 FT CYTOPLASMIC (POTENTIAL).
 FT IG-LIKE C2-TYPE 1.
 FT IG-LIKE C2-TYPE 2.
 FT IG-LIKE C2-TYPE 3.
 FT IG-LIKE C2-TYPE 4.
 FT IG-LIKE C2-TYPE 5.
 FT PROTEIN KINASE.
 FT ATP (BY SIMILARITY).
 FT NP BIND 596
 FT BINDING 616

DR InterPro: IPR000719; Prot_kinase.
 DR InterPro: IPR001824; Rtkinaseiii.
 DR Pfam: PF00047; ig; 3.
 DR Pfam: PF00069; pkinase; 1.
 DR ProDom: PD000001; Prot_kinase; 2.
 DR SMART: SM00408; IGC2; 1.
 DR SMART: SM00219; TyrKc; 1.
 DR PROSITE: PS00835; IG LIKE; 1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE: PS00111; PROTEIN_KINASE_DOM; 1.
 DR PROSITE: PS00109; PROTEIN_KINASE_TMR; 1.
 DR PROSITE: PS00240; RECEPTOR_TYR_KIN_III; 1.
 KW Proto-oncogene; Tyrosine-protein Kinase; Receptor; Transmembrane;
 KW Transferase; Glycoprotein; Phosphorylation; ATP-binding; Signal;
 KW Immunoglobulin domain.
 FT SIGNAL 1 24
 FT CHAIN 25 960
 FT DOMAIN 25 505
 FT TRANSMEM 506 530
 FT DOMAIN 531 960
 FT DOMAIN 575 913
 FT NP_BIND 581 589
 FT BINDING 609 609
 FT ACT_SITE 777 777
 FT MOD_RES 808 808
 FT CARBOHYD 76 76
 FT CARBOHYD 135 135
 FT CARBOHYD 149 149
 FT CARBOHYD 269 269
 FT CARBOHYD 286 286
 FT CARBOHYD 306 306
 FT CARBOHYD 318 318
 FT CARBOHYD 338 338
 FT CARBOHYD 343 343
 FT CARBOHYD 356 356
 FT CARBOHYD 453 453
 FT CARBOHYD 469 469
 SQ SEQUENCE 960 AA; 107311 MW; 0893850527AB6F6 CRC64;
 Query Match 23.0%; Score 1215; DB 1; Length 960;
 Best Local Similarity 32.3%; Pred. No. 1.4e-75;
 Matches 314; Conservative 165; Mismatches 370; Indels 124; Gaps 34;
 QY 66 ALRPOSSGTVYAAAEVDVDSITQLVLDAPGNISCLVWFKHSS-----LNCOPHPD 119
 DB 18 SLIPAGGVPHESSILVNVKGBELRLKNEGP-----VTWNQNSDPSAKTRISNEKEWH 73
 QY 120 LQNRGV--VSMVILKMTQAGEYLLFIOSEATNYTLFTVSIKNTLLYTLRRPYPRKME 177
 DB 74 TKNATIRDIGRECKSGKSVNSFYVFKDP---NVLELV---DSLIIY-----GKED 119
 QY 178 NODALVCISSEVPEP-IVEMVLCDGSGESCKEESPAVVKKE-----KVLHELFGTDIRCC 232
 DB 120 SDILLVC---PLTDPDLNFTLRKDGKPLKPNMTFIPNQKGIILKVNQSFKGCYQCL 176
 QY 233 AR-NELGRECTRLFTIDLNQTPL-OTLPLQ-----FLKVGEBPLMRCKAVVNHGFGCL 283
 DB 177 AKHNGVEKISEHF---LNVRPVHKALPVITLSKVELLKEGEFVETCIITDVSUVKA 233
 QY 284 TWELNKALEEGNYFMSTYSTNRNTRMIRILFAPVSSVARNDGYTCSSKHP--SQSAL 341
 DB 234 SWISYKSAIVTSKRNLDGYVERK---LTNIRSGVNDSGEFTC-QAENPFPGKTNAT 288
 QY 342 VTI--VGKGFN---ATNSDEYIDQYEEFCFSPKAYPQIR-CTWTFSRKSPFCEQK 395
 DB 289 VTLKALAKGFVLFATMTNTTIDINAGQNG--LTVEYEAYPKPEKVVMMYMETL---QN 343
 QY 396 GLDNGYSISKFNHKK-----OPGEYIFHAENDDAQFTKMTFLNTRRKPOVLA 443
 DB 344 SSDHYVKFTVGNNSYSELHLTRLKGTGGIYTFVVSNDASSSTVFNVYVTKKEIILT 403
 QY 444 EASASQ--ASCFSGDGYPLFSWTWKCKSDKSPNCTEITEGVNWRKANRVFGQWSSSTL 501
 Db 404 LDMLGNDILQCVATGFPAPTYIYWFPCPTQECRLDSPISPMVKVS-----YINSSYP 457
 QY 502 NMSEAIGKFLYKCCAYNSLGTSCETILLNSPGPPFF-----IODNISFYA-----TIGV 550
 Db 458 SPERILVESTVNASMFKSTGTICCEASSNGDKSVFFNFPAIKEQIRTHRTFLPLIAFGV 517
 QY 551 CLLFIVVLTLLICHKKQFYESQLQMV-QVTGSSDNEYFYVDREYEDLAKWEPREN 609
 Db 518 AAGLMCIIVMILVIYLOKPKYEQVMKVEEING--NNYVVIDPTQLPYDHKWEPFNR 574
 QY 610 LEFGKVLGSGAGFKVMNATAYGISKTGVSIOVAVKMLKEKADSSREALMSLKMWTOLG 669
 Db 575 LSFGLTKGAGAFKGVVEATATGLFKSDAAMTVAVKMLKPSAHLTEREALMSLKVLSYLG 634
 QY 670 SHENIVNLLGACTLSGPIYLIFEYCCYGDLLNLYRSKREKF-----HRTWTEIFKEHNF 725
 Db 635 NHINIVNLLGACTIGGPTLVITEYCCYGDLLNFLARKRDSFICPKHEEHAAYVENL-- 692
 QY 726 YPTFQSHNS-----SMPGSRVQIHPDSQISGLHNSFHSDEIYENQKLEEBE 778
 Db 693 --LHQAEPTADAVNEYMDKPGVSYAVPPKADKKRPVKSGSYTDQD---VTLSLMEDDE 746
 QY 779 DLNVLTPEDLICFAYQVAKGMEFLFKSCVHEDLAARNVLVTHGKVKVICDFGLARDIMS 838
 Db 747 --LALDVEDLLSFSYQVAKGMSFLASKNCIHRDLAARNILLTHGRITKICDFGLARDIRN 804
 QY 839 DSNVYVRGNARLPVKWMAPESEFEGIYITKSDVWSYGILLWEIFSLGVNPPYPCIPVDANF 898
 Db 805 DSNVYVRGNARLPVKWMAPESEFENCVYTFESDVWSYGILLWELFSLGSSPYGMPVDSKF 864
 QY 899 YKLQNGFMQOPFVATBEIYIIMOSWAFDSKRKPSPNLTSFLGCGQLADAEAEAMQNV 958
 Db 865 YKMIKEGYRMFSPESCPPEMYDIMKSCWDADPLQPTFKQIVQLIEQQLSDNAPRYAN- 923
 QY 959 DGRVSECPHTYQN 971
 Db 924 ---FSTPPSTQGN 933
 RESULT 13
 KPMS MOUSE
 ID KPMS_MOUSE STANDARD; PRT; 977 AA.
 AC P09581; Q9DBH9;
 DT 01-MAR-1989 (Rel. 10, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Macrophage colony stimulating factor I receptor precursor (CSF-1-R)
 DE (BC 2.7.1.112) (fms proto-oncogene) (c-fms).
 GN CSF1R OR CSFMR OR FMS;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=8217329; PubMed=2966922;
 RA Rothwell V.M., Rohrschneider L.R.;
 RT "Murine c-fms cDNA: cloning, sequence analysis and retroviral
 RT expression.";
 RL Oncogene Res. 1:311-324 (1987).
 RN [2]
 RP REVISIONS.
 RA Rothwell V.M.;
 RN [3]
 RL Submitted (SEP-1988) to the EMBL/GenBank/DBJ databases.
 RX MEDLINE=93181280; PubMed=8441691;
 RA de Parseval N., Boredeaux D., Gisselbrecht S., Sola B.;
 RT "Reassessment of the murine c-fms proto-oncogene sequence.";
 RL Nucleic Acids Res. 21:750-750(1993).
 RN [4]
 RP SEQUENCE FROM N.A.

Db 88 EDPMAGSTTHLYVVKDPAHSMNLLAQEVTV-----VEGQEAIV----- 124
Qy 189 VPEIVIEWLVDQGESCKEESPAVVK-----EEKVLHFLGTDIRCCAR 234
Db 125 LPCLITDPAKDSV-SLMREGGRVLRKTVTFSSAWRGFIIRKAKVL-----DSNTYVCKT 179
Qy 225 NELGRECTL-FTIDLN-----QTPQTL-PQLFLKV-GEPLMTRCKAVHNVHFGTLTWEL 287
Db 180 MYNGESTSTGWLKNVRVHPFPQIKLEPSKLVIRGEAAQIVCSATNAEVGNVILKR 239
Qy 288 ENKALE-----EGNYFEMSTYSTNRTMIRILFAFVSSVARNDTYTC-SSSKHPSQS 339
Db 240 GDTKLEIPLNSPDQNYK-----KVRALSINAVFODAGIYCSVNSDVGT 288
Qy 340 ALVT--IVKGFINATN--SSEYEDIDOEYEFCEPFRKAYPOIR-CTWTFSRKSPCEOK 395
Db 289 ATMFQVVESAYLNTSEQSLQEVSGDSLITVHADAYPSIOHYNWYLYGPFEE-DQR 347
Qy 396 GLD-----NGYSISKFNH--KHQPGYIPIHAENDDAQFTQWFTLNIRRKPOVLAEAS 446
Db 348 KLEFITQRAIYRYTKLFLNRVKASEAGQVFLMAQNKAGNNLTFTLTLYPEVSVTW 407
Qy 447 ASQAS-----CFSDGYPLPWTWKCKDKSPNCTEITEGVN-----RKANRKVFGQWSS 498
Db 408 PVNGSDVLCFDSVGYDQPSVYTWECRGHTDRCDQAQALQVNDTHPEVLSQKPFVKV 467
Qy 499 STLNSEAIKGLVKCAYNLSGTSCETILLNSPGPPFPFIQDNISFYATIGVC---LLFI 555
Db 468 SOLPIGTLKHNMTYFCKTHNSVGNSSQYFRAVSLGQSKQLPDESFTPPVAVCMVMSLL 527
Qy 556 VVLTLLIHKYKQFYESQLOMQVGTSSDNEFYVYDFREYDYDKWEPPEPNELEFGKV 615
Db 528 VLLLLLLLYKQKPYQVWKIIE--RYEGNSYTFIDPTQLPYNKWKPEPNNLQFGKT 585
Qy 616 LSGAFGKVMNATYAGISKTGVSIOAVKMLKEKADSSEREALMSLKMWTOLGSHENIV 675
Db 586 LGAGAFKQVVEATFGLGKEDAVLKVAVKMLKSTAHADKEALMSLKMILGQHENIV 645
Qy 676 NLLGACTSLGPIYLFYCCYCDLLNLYLSKRE-----KPHRTWT 715
Db 646 NLLGACTHGGPVLVITEYCCYCDLLNLYLSKRE-----KPHRTWT 715
Qy 716 EYFKEHNFQYPTFQSHPNMSPGSEVQIHDPDSQISGLHGNFSHSEBIEYENOKRLE 775
Db 706 YVRRDSGFS-----SQVDYVEMRPVSTSSDSFPKQD-LDKEPSRPLE 749
Qy 776 EEDLNVLTFEDLLCFAYQVAKMEFLKSCVHRDLAARNVLVTHGKVKIKCDFGLARD 835
Db 750 -----LWDLHFSQVQAQMAFLAKNCIHRDVAARNVLTSGHVAKIGDFGLARD 800
Qy 836 IMSDSNVYVVRGNARLPVKWAPSLFEGYITIKSDVMSYGILLWEIFSLGVNYPGPVVD 895
Db 801 IMWDSNVYVVRGNARLPVKWAPSLFEGYITIKSDVMSYGILLWEIFSLGVNYPGPVVD 895
Qy 896 ANFYKLIQNGFKMDQPFYATEEYIIMQSCWAFDSRKRPSFPNLTFLGCOLADAEAMY 955
Db 861 NKFKYKLVKDGQYMAQVPFAPKNIYSIMQSCWDLPTRRPTFOICELL-----QEOARL 914
Qy 956 QNVGDRVSECP 966
Db 915 ERRDQYANLP 925

RESULT 15

PGDS RAT STANDARD; PRT; 1088 AA.

AC P20786;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Alpha platelet-derived growth factor precursor (EC 2.7.1.112)
DN (PDGF-R-alpha).
GN PDGFRA.

OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley;
RX MEDLINE=90220609; PubMed=2157969;
RT Lee K.H., Bowen-Pope D.F., Reed R.R.;
RT "Isolation and characterization of the alpha platelet-derived growth
RT factor receptor from rat olfactory epithelium.";
RL Mol. Cell. Biol. 10:2237-2246(1990).
RN (2)
RP SEQUENCE OF 33-524 FROM N.A.
RX MEDLINE=93305723; PubMed=8318539;
RA Herren B., Weyer K.A., Rouge M., Loetscher P., Pech M.;
RA "Conservation in sequence and affinity of human and rodent PDGF
RT ligands and receptors.";
RL Biochim. Biophys. Acta 1173:294-302(1993).
CC -!- FUNCTION: THIS RECEPTOR BINDS PLATELET-DERIVED GROWTH FACTOR AND
CC HAS A TYROSINE-PROTEIN KINASE ACTIVITY. THIS RECEPTOR CAN BIND
CC EITHER PDGF-A OR PDGF-B.
CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC tyrosine phosphate.
CC -!- SUBUNIT: DIMER OF EITHER ALPHA-ALPHA, BETA-BETA OR ALPHA-BETA
CC SUBUNITS.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-
CC PROTEIN KINASES.
CC -!- SIMILARITY: Contains 5 immunoglobulin-like C2-type domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
DR EMBL; M63837; AAA40743.1; ALT_INIT.
DR EMBL; Z14118; CAA78488.1; -.
DR PIR; A34710; PERTGA.
DR HSP; P11362; LFGK.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR001824; RTKinaseIII.
DR InterPro; IPR001245; Tyr_kinase.
DR Pfam; PF00047; Ig; 4.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 2.
DR SMART; SM00409; IG_3.
DR SMART; SM00219; TyrcK; 1.
DR PROSITE; PS00835; IG_LIKE; 3.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE; PS00240; RECEPTOR_TYR_KIN_III; 1.
KW Tyrosine-protein kinase; Receptor; Transmembrane; Glycoprotein;
KW Transferase; Phosphorylation; ATP-binding; Immunoglobulin domain;
KW Signal; Repeat.
FT CHAIN 1 23 ALPHA PLATELET-DERIVED GROWTH FACTOR
FT RECEPTOR.
FT DOMAIN 24 524 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 525 548 POTENTIAL.
FT DOMAIN 549 1088 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 25 112 IG-LIKE C2-TYPE 1.
FT DOMAIN 213 305 IG-LIKE C2-TYPE 3.
FT DOMAIN 318 409 IG-LIKE C2-TYPE 4.
FT DOMAIN 592 953 PROTEIN KINASE.
FT NP_BIND 598 606 ATP (BY SIMILARITY).

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model
Run on: August 26, 2003, 07:20:28 ; Search time 83.5421 Seconds
(without alignments)
3067.272 Million cell updates/sec

Title: US-09-919-408A-4
Perfect score: 5274
Sequence: 1 MPALARDAGTVPLLVVFSAM.....PFSREMDGLLSPOAQVEDS 993

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues
Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : SPTREMBL 23:*
- 1: sp_archaea:*
 - 2: sp_bacteria:*
 - 3: sp_fungi:*
 - 4: sp_human:*
 - 5: sp_invertebrate:*
 - 6: sp_mammal:*
 - 7: sp_mbc:*
 - 8: sp_organelle:*
 - 9: sp_phage:*
 - 10: sp_plant:*
 - 11: sp_rodent:*
 - 12: sp_virus:*
 - 13: sp_vertebrate:*
 - 14: sp_unclassified:*
 - 15: sp_rvirus:*
 - 16: sp_bacteriap:*
 - 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1707.5	32.4	406	11 Q8BR20	Q8br20 mus musculus
2	1272.5	24.1	976	13 Q8JPR5	Q8jfr5 brachydanio
3	1271.5	24.1	976	13 Q9W755	Q9w755 brachydanio
4	1271	24.1	979	11 Q8C8K9	Q8c8k9 mus musculus
5	1265.5	24.0	978	11 Q63702	Q63702 rattus ratt
6	1263.5	24.0	978	11 Q63116	Q63116 rattus norv
7	1262.5	23.9	964	6 Q97744	Q97744 sus scrofa
8	1262	23.9	984	13 Q8AXC6	Q8axc6 fugu rubrip
9	1261.5	23.9	964	6 Q9TQ01	Q9tcq1 sus scrofa
10	1261.5	23.9	964	6 Q9TQ00	Q9tcq0 sus scrofa
11	1260.5	23.7	979	6 Q8WN23	Q8wn23 canis famli
12	1247	23.6	972	4 Q99662	Q99662 homo sapien
13	1244	23.6	978	6 Q9XS93	Q9xs93 canis famli
14	1230.5	23.3	977	13 Q98SU1	Q98sul danio nigro
15	1229	23.3	945	6 Q77589	Q77589 equus cabal
16	1228	23.3	974	13 Q98SU3	Q98su3 danio dangi

17	1226	23.2	954	13 Q91909	Q91909 xenopus lae
18	1223.5	23.2	977	13 Q918N6	Q918n6 brachydanio
19	1221	23.2	948	6 Q9TDT7	Q9tdt7 trichosurus
20	1219	23.1	724	6 Q9WYN0	Q9wyn0 bos taurus
21	1214	23.0	977	13 Q98SU2	Q98su2 danio ketti
22	1213	23.0	975	13 Q98SU4	Q98su4 danio albol
23	1211	23.0	977	13 Q98SU0	Q98su0 oncorhynch
24	1192.5	22.6	992	13 Q8AXU0	Q8axu0 fugu rubrip
25	1163	22.1	1062	13 Q8AXC7	Q8axc7 fugu rubrip
26	1162	22.0	1078	13 Q8AXC8	Q8axc8 fugu rubrip
27	1141.5	21.6	1059	13 Q9DE49	Q9de49 brachydanio
28	1126	21.4	1087	13 Q9PUF6	Q9puf6 gallus gall
29	1111.5	21.1	1097	11 Q8R406	Q8r406 rattus norv
30	1087.5	20.6	1106	4 Q8N5L4	Q8n5l4 homo sapien
31	1080	20.5	1048	13 P79749	P79749 fugu rubrip
32	1079	20.5	1019	13 Q8UVR8	Q8uvr8 fugu rubrip
33	1046	19.8	923	6 Q97745	Q97745 sus scrofa
34	1038	19.7	986	13 Q8UVR9	Q8uvr9 fugu rubrip
35	965.5	18.3	563	11 Q925F7	Q925f7 rattus norv
36	954.5	18.1	1302	13 Q8AXB3	Q8axb3 brachydanio
37	952.5	18.1	1173	13 Q9PTL0	Q9ptl0 brachydanio
38	950.5	18.0	1301	13 Q8UWU9	Q8uwu9 brachydanio
39	948.5	18.0	1327	13 Q8QHL3	Q8qhl3 gallus gall
40	934	17.7	1345	11 Q8VCD0	Q8vcd0 mus musculu
41	932	17.7	1379	13 P79701	P79701 coturnix co
42	916.5	17.4	323	11 Q9EQ22	Q9eq22 rattus norv
43	909.5	17.2	1363	11 Q91ZT1	Q91zt1 rattus norv
44	892.5	16.9	323	11 Q9EQ24	Q9eq24 rattus norv
45	841	15.9	824	13 Q90749	Q90749 gallus gall

ALIGNMENTS

RESULT 1
Q8BR20
ID Q8BR20 PRELIMINARY; PRT; 406 AA.
AC Q8BR20;
DT 01-MAR-2003 (TRENBLrel. 23, Created)
DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE FMS-like tyrosine kinase 3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Brain;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
60,770 full-length cDNAs."
RL Nature 420:563-573 (2002).
DR EMBL; AK045865; BAC32514.1; -.
SQ SEQUENCE 406 AA; 45834 MW; 155394D167D7D37 CRC64;

Query Match	32.4%;	Score 1707.5;	DB 11;	Length 406;
Best Local Similarity	80.0%;	Pred. No. 1.3e-135;		
Matches 324;	Conservative 30;	Mismatches 50;	Indels 1;	Gaps 1;
Qy	1	MPALA-RDAGTVPLLVVFSAMIFGTTNQDLPVIKCVLINHKNDSSVGKSSSYPMVSES	59	
Db	1	MRALQSRDRELLLVLSVILEVTNQLPVIKCVLISHENNGSSAGKSSYRMVRS	60	
Qy	60	PEDLGCAIRPOSSGTVEAAAVEVDVNSAITLQVLVDPAGNISCLWVPKHSINCOHPHD	119	
Db	61	PEDLQCAPRROSEGTVEAAATVEAEGSITLQVQLATPGDLSCLVWPKHSILGQCPHFD	120	
Qy	120	LQNRGVSWLTKMTQEQAYLLFIQSEANTYTLFTVSIKNTLLYTLRRPYFKMENO	179	
Db	121	LQNRGVSWLTKMTQEQAYLLFIQSEANTYTLFTVSIKNTLLYTLRRPYFKMENO	180	

Qy 180 DALVCISEVPEPIVWVLCDSQGESCKEESPAPVVKKEKVLHELFGTDIRCCARNELGR 239
 Db 181 DALLCISGVEPTVWVLCSSHRESCKEESPAPVVRKEKVLHELFGTDIRCCARNALGR 240
 Qy 240 ECTRLFTIDLANOTPTLLPQLFLKVGEPILWIRCKAVHVNHGFLTWELNKALEEGNYFE 299
 Db 241 ECTRLFTIDLANOQFSTLPQLFLKVGEPILWIRCKAVHVNHGFLTWELNKALEEGSYFE 300
 Qy 300 MSTYSTNRTMIRILFAFVSSVARNDTGYTSSSKHPSQSALVTIVGVGINATNSSBDY 359
 Db 301 MSTYSTNRTMIRILFAFVSSVGRNDTGYTSSSKHPSQSALVTILEKGINATSSQBEY 360
 Qy 360 BIDYBEFCFVRKAYPQIRCTWTFSRKSPCEQKGLDNGYSIS 404
 Db 361 BIDPYEKFCFVRKAYPQIRCTWTFSSQSPCEQKGLDNGYRLS 405

RESULT 2
 Q8JFR5 PRELIMINARY; PRT; 976 AA.
 AC Q8JFR5;
 DT 01-OCT-2002 (TRENBLrel. 22, Created)
 DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)
 DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
 DE SI:did116L04.1 (kit receptor (Tyrosine kinase)).
 GN KIT.
 OS Brachydanio rerio (Zebrafish) (Danio rerio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Danio.
 OC NCBI_TaxID=7955;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Pandian R.;
 RL Submitted (AUG-2002) to the EMBL/GenBank/DBSJ databases.
 DR EMBL; AL691516; CAD3458.1; -;
 DR InterPro; IPR003599; IG.
 DR InterPro; IPR007110; IG-like.
 DR InterPro; IPR003598; IG c2.
 DR InterPro; IPR003006; IG_MHC.
 DR InterPro; IPR000719; Prot_kinase.
 DR InterPro; IPR01824; RtkinaseIII.
 DR InterPro; IPR002290; Ser_thr_kinase.
 DR InterPro; IPR001245; Tyr_kinase.
 DR Pfam; PF00047; Ig_3.
 DR Pfam; PF00069; pkinase; 1.
 DR Pfam; PF00001; Prot_kinase; 2.
 DR SMART; SM00409; IG; 4.
 DR SMART; SM00408; IGC2; 3.
 DR SMART; SM00220; S_TK; 1.
 DR SMART; SM00219; TyKc; 1.
 DR PROSITE; PS00835; IG LIKE; 2.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS00011; PROTEIN_KINASE_TM; 1.
 DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
 DR PROSITE; PS00240; RECEPTOR_TYR_KIN_III; 1.
 KW ATP-binding; Immunoglobulin domain; Kinase; Receptor; Transferase.
 SQ SEQUENCE 976 AA; 109293 MW; A56921FA48DE58ED CRC64;

Query Match 24.1%; Score 1272.5; DB 13; Length 976;
 Best Local Similarity 38.5%; Pred. No. 2.8e-98;
 Matches 294; Conservative 122; Mismatches 262; Indels 85; Gaps 20;

Qy 253 PQTTLPO---LPLKVGEPILWIRCKAVHVNHGFLTWELNKALEEGNYFEMSTYSIN 306
 Db 206 PPITGQFORVLTQGEKLSLCSSTNSVNDIAVWKAENGVPVHQNSHLTTPITHV 265
 Qy 307 RTMIRILFAFVSSVARNDTGYTSSSKHPSQSA---LVTIVGKGINATN-SSDYEID 362
 Db 266 RTAILSL-----SSVTQDAGNVSCEAINEKGTAKPWNVIYKGPINITSVDNSTRRV 321
 Qy 363 QYEEFCFVRKAYPQIRCTWTFSRKSPCEQKGLDNGYSISKFNH----- 409

Db 322 AGESLSLRVWNAFKPHTFSWSYS-----GVKLTTNTDVIHTRTHGNSVT 368
 Qy 410 -----KHQPEYIFHAENDDAQTKMTLIRKPKQVLAESA--SQASCFSDGYP 458
 Db 369 SELKVLRLKVSSEGIYTFSCILNRDATIRQTEPVHVISKPQIVSYEGPIDGQVRCVAEGYP 428
 Qy 459 LPSWTW-----KKCSDKSPNCTEETETEGVWNRKANRKFVGQWVSSSTLANSEAIKGLF 511
 Db 429 TPQIKWYCDLPHSRCNLL-NATQE-BEDVVTITMTNPPFGKAVESRLNITK--NNYA 484
 Qy 512 VKCAYNSLGRSCETILLNSPGPP---FIQDNISFVATIGVCLLFIIVLTILLICHKYKK 568
 Db 485 TLECVASANGEIVTLPSISENTVPHELFTPLLIGFVAAVILVILLIVLT-----YKMQ 540
 Qy 569 QFRYEQLMQVQVQVSSDNEYFYVDREYEDLKWEPPREMLEFKVLGSGAFGKVMNAT 628
 Db 541 KPQIQIQWKVIE--GIHGNVYVIDPTQLPYDQWEPFRDKLRFGLTGLSGAFGKVEAT 598
 Qy 629 AYGISKTVSIOVAVKMLKEKADSSEREAALSELKMTQLGSHENIVNLLGACTILSGPY 688
 Db 599 AYGMSKADVTMTAVKMLKPSAHATEKALSELKVLVSLGNHINIVNLLGACTVGGPTL 658
 Qy 689 LIPEYCCVGLLNLVLRSKREKFRHTWTEIFKEHNFSFYPTFQSHNSNMPGSREYQ--- 744
 Db 659 VITEYCCFGDLNLFERRRVFYVT---TLGEDAYYRNVMQSEPDNRNGYMTWKPSVL 715
 Qy 745 -IHPDSQISGLHGNSHSEDEIEYENQKRLEEDLNVLTFEDLLCFAYQVAKGMFELE 803
 Db 716 GILSENRRSLKNGDSYSDAV-----SEILOEDGLTLDT-EDLLSFSYQVAKGMDFLA 769
 Qy 804 FKSCVHRDLAARNVLVTHGKVKVICDFGLARDINSDSNVYVGNARLPVKWMAPESLPEG 863
 Db 770 SKNCIHRDLAARNILLTQGRVAKICDFGLARDITDTSNYYVKGARLPVKWMSPESEIFEC 829
 Qy 864 IYTIKSDVWSYGIILLWEIFSLGVNPPGIPVDANFYKLIQNGFKMDQDPYATEEVIING 923
 Db 830 VITFESDVMWSYGIILLWEIFSLGVNPPGIPVDANFYKLIQNGFKMDQDPYATEEVIING 889
 Qy 924 SCWAFDSRKRFSPNLTSLFGCLQADAEAMYQNVGDRVSECP 966
 Db 890 SCWDADPVKRFSPFSKIVEKIQISDSTKHYLVNFSRLPAAP 932

RESULT 3
 Q9W755 PRELIMINARY; PRT; 976 AA.
 AC Q9W755;
 DT 01-NOV-1999 (TRENBLrel. 12, Created)
 DT 01-NOV-1999 (TRENBLrel. 12, Last sequence update)
 DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
 DE Kit receptor tyrosine kinase.
 GN KIT.
 OS Brachydanio rerio (Zebrafish) (Danio rerio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Danio.
 OC NCBI_TaxID=7955;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99396707; PubMed=10393121;
 RA Parichy D.M., Rawls J.F., Pratt S.J., Whitfield T.T., Johnson S.L.;
 RT "Zebrafish sparse corresponds to an orthologue of c-kit and is
 RT required for the morphogenesis of a subpopulation of melanocytes, but
 RT is not essential for hematopoiesis or primordial germ cell
 RT development.";
 RL Development 126:3425-3436(1999).
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-
 CC PROTEIN KINASES.
 DR EMBL; AF153446; AAD41890.1; -;
 DR HSSP; P11362; 1FGK.
 DR ZFIN; ZDB-GENE-980526-464; kit.

RESULT 6			
Q63116	PRELIMINARY;	PRT;	978 AA.
AC	Q63116		
DT	01-NOV-1996 (TREMELrel. 01, Created)		
DT	01-NOV-1996 (TREMELrel. 01, Last sequence update)		
DT	01-MAR-2003 (TREMELrel. 23, Last annotation update)		
DE	C-kit receptor tyrosine kinase.		
OS	Rattus norvegicus (Rat).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.		
OX	NCBI_TaxID=10116;		
EN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=Sprague-Dawley;		
RX	MEDLINE=92003944; PubMed=1912577;		
RA	Tsujimura T., Hirota S., Nomura S., Niwa Y., Yamazaki M., Tono T.,		
RA	Mori E., Kim H., Kondo K., Nishimune Y., Kitamura Y.;		
RT	"Characterization of a mutant allele of c-kit gene."		
RT	tyrosine kinase domain of c-kit gene."		
RL	Blood 78:1942-1946 (1991).		
CC	!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).		
CC	!- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-		
CC	PROTEIN KINASES.		
DR	EMBL; D12524; BA02094.1; -		
DR	HSSP; F11362; IFGK.		
DR	InterPro; IPR007110; Ig-like.		
DR	InterPro; IPR003598; Ig C2.		
DR	InterPro; IPR003006; Ig MHC.		
DR	InterPro; IPR000719; Prot kinase.		
DR	InterPro; IPR001824; RTKinaseIII.		
DR	InterPro; IPR001245; Tyr_kinase.		
DR	Pfam; PF00047; Ig; 1.		
DR	Pfam; PF00069; pkinase; 1.		
DR	ProDom; PD000001; Prot kinase; 2.		
DR	SMART; SM00408; IGC2; 1.		
DR	SMART; SM00219; TyrKc; 1.		
DR	PROSITE; PS00835; IG LIKE; 3.		
DR	PROSITE; PS00107; PROTEIN KINASE ATP; 1.		
DR	PROSITE; PS00111; PROTEIN KINASE DOM; 1.		
DR	PROSITE; PS00109; PROTEIN KINASE TYR; 1.		
DR	PROSITE; PS00240; RECEPTOR_TYR_KIN_III; 1.		
KW	ATP-binding; Glycoprotein; Immunoglobulin domain; Kinase;		
KW	Phosphorylation; Receptor; Transferase; Transmembrane;		
KW	Tyrosine-protein kinase.		
SQ	SEQUENCE 978 AA; 109341 MW; 0958333F19889051 CRC64;		
Query Match 24.0%; Score 1263.5; DB 11; Length 978;			
Best Local Similarity 33.8%; Pred. No. 1.6e-97;			
Matches 334; Conservative 171; Mismatches 361; Indels 121; Gaps 37;			
QY	35	CYLLINHNKNDSSVGKSSYPWVSESPELGLCALRPOSSGTVVEAAAVEVDVSASITLQVL 94	
Db	12	CVLLVLLRQGTQSQSPASGPEPSPP-----SIQPAQS-----ELIVEAGDTIRLT 57	
QY	95	VDAPGNISCLVFKHSLNCQPHFDLQNRGVSMVLKMTL-TOAGEYLLFIQSEATNYT 153	
Db	58	CTDPAPVK--WTFE-----ILDVRIENKQSEWIREKAEATHGKYTCVSGSLRSSI 107	
QY	154	ILFTVSRN-TLLYTRRPFYRKQENQDALVCISSEVPEIWEVLCDQSGESCKEESPA 212	
Db	108	YVF---VRDPAVLFLVGLFLFGK-EDNDALVRCPFLDPQ-VSNYSLETCDGKSLPTDLKF 162	
QY	213	V-----VKKEKVLHFLPGDTDIRCCARNELGRETFLFTD-----LNQTPQTTLPOL 260	
Db	163	VNPKAGITIKNRYAHRLC-----IKCAQREGKWRSDKFTLUKRAAIKAIIPVSPET 219	
QY	261	--FLKVGPELWIRKAVHVNHGFLTW-----ELENKALEEGNYFEMS--TYSTNRTMIR 311	
Db	220	SHLLKEGDTFTVICTIKDVSVDVSMWIKLNPQSQKAVKRNHSGQDFNRYEQETLT- 278	
QY	312	ILFAFVSVARNDTGYTTCSSSKHPQSALVT---IVGKGFN---ATNSSEYEDIOYE 365	

RESULT 7			
ID	O97744	PRELIMINARY;	PRT; 964 AA.
AC	O97744;		
DT	01-MAY-1999 (TREMELrel. 10, Created)		
DT	01-MAY-1999 (TREMELrel. 10, Last sequence update)		
DT	01-MAR-2003 (TREMELrel. 23, Last annotation update)		
DE	Mast/stem cell growth factor receptor (Fragment).		
GN	KIT1*0101.		
OS	Sus scrofa (Pig).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.		
OX	NCBI_TaxID=9823;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=Hampshire breed; TISSUE=Leukocyte;		
RX	MEDLINE=98391767; PubMed=9724328;		
RA	Marklund S., Kijas J., Rodriguez-Martinez H., Roennstrand L., Funa K.,		
RA	Moller M., Edfors-Lilja I., Andersson L.;		
RT	"Molecular evolution of the dominant white phenotype in the domestic pig."		
RL	Genome Res. 8:826-833 (1998).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=Hampshire breed; TISSUE=Leukocyte;		
RA	Marklund S.;		

Db	279	-----ISSARVNDSGVFMCYANNTFGSANVTTLTKVVEKGFNIFPVKNT--VFVTDGE 331	
QY	366	EFCSVRFKAYPO-IRCTWTFRSKSPCBQKGLD-----NGYISKFCNKH-----OP 413	
Db	332	NVDLVFEAYPKPEHQIYINRT--PTNREGDYVSDNQSNIRVYNELRLTRKGTGEC 389	
QY	414	GEYIFHAENDDAQFTMTINIRRPQVLA--EASASQSCFSDGYPLPSWTWKKCDKS 471	
Db	390	GYTTLVNSDVASVTPDVYVNTPEILTYDRLMNGRLQCVAAGPPEPTIDWYFCTGAE 449	
QY	472	PNCETBEITE-GVNNRKANKRVGQWSSSTLNMSEAIGFLVKCCAYNSLGTSCETILLN 530	
Db	450	QRCTVPPVDVQIQNASVSPFGKLVQSSIDSSVFRHNGTVECKASNAVGK--SAFEN 507	
QY	531	SPGPPEP-----IQDN-----ISFYATIGVCLLFIWLTLLCHKYKGFYRESOL 576	
Db	508	-----FAFKGNSKEIQPHLFTPLLLIGFVVVTA--LMGIIVWL--AYKLPQPMTEVQW 559	
QY	577	QMV-QVTSSDNEYFVDFREYEDLKWFFRENLEFFGVLGSGAGFKVMNATAYGISKT 635	
Db	560	KVVEING---NNYVYIDPTQLPYDHKWEFPNRNLSFGKTLGAGAFGVVEATAYGLIKS 616	
QY	636	GVSIQAVVMLKEKADSSREALMSLKMVTLGSHENIVNLGACTLGSPIYLIEYCC 695	
Db	617	DAAMTVAVKMLKPSAHLTEREALMSLKVLSYLGNNHNIIVNLGACTVGPGTLVITEYCC 676	
QY	696	YGDLLNLSKREKFRHTWTEIFKEHNFYTFQSHPNSSMPGSR-VQIHDPDSOISG 754	
Db	677	YGDLLNFLRRKDSF--IFSQEQADAALYKMLLSKSSCDSSNEYMDMKPGSVYVP 734	
QY	755	LHGNSFHS---EDELEYENQKELEBEDLNLTFFEDLLCFAYOVAKGMELEFKSCVHRD 811	
Db	735	TKTKRRSARIDSYIERDVTVAIMEDDEL-ALDELLSFSYOVAKGMAFLASKNCIHRD 793	
QY	812	LAARNVLVTHGKVKICDFGLARDIMSDSNVYVVRGNARLPVKWMAPELFEGIYTIKSDV 871	
Db	794	LAARNILLTHGTRITKICDFGLARDIRNDSYVYVVGARLPVKWMAPELFEIENCVTPESDV 853	
QY	872	WSYGILLWEIPSLGVNYPGIPVDANFYKIQNGFQMDQPFYATEYIIYMQSCWAFDSR 931	
Db	854	WSYGIFLWELFSLGSSPYGMPVDSPFYKMKIEGFRMLSPHAPAAAMYEMKTCWDADPL 913	
QY	932	KRPSPNLTSFLGCLADAAEAMYNQV 958	
Db	914	KRPTFKQVVQLIEKQISDSSKHIYSNL 940	

RL Thesis (1997), Department of Animal Breeding and Genetics.
 RL Swedish University of Agricultural Sciences, Uppsala, SWEDEN.
 CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-
 CC PROTEIN KINASES.
 DR EMBL: AJ223228; CAAL1196.1; -.
 DR HSSP: P11362; IFGK.
 DR InterPro: IPR003599; Ig.
 DR InterPro: IPR007110; Ig-like.
 DR InterPro: IPR003006; Ig_MHC.
 DR InterPro: IPR000719; Prot_kinase.
 DR InterPro: IPR001824; RTK_kinaseIII.
 DR InterPro: IPR001245; Tyr_kinase.
 DR Pfam: PF00047; Ig; 1.
 DR Pfam: PF00069; pkinase; 1.
 DR ProDom: PD000001; Prot_kinase; 2.
 DR SMART: SM00409; IG; 3_kinase; 2.
 DR SMART: SM00219; TyrcK; 1.
 DR PROSITE; PS00835; IG_LIKE; 3.
 DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
 DR PROSITE; PS00111; PROTEIN KINASE DOM; 1.
 DR PROSITE; PS00109; PROTEIN KINASE_TYR; 1.
 DR PROSITE; PS00240; RECEPTOR_TYR_KIN_III; 1.
 DR ATP-binding; Glycoprotein; Immunoglobulin domain; Kinase;
 KW Phosphorylation; Receptor; Transmembrane;
 KW Tyrosine-protein kinase.
 FT NON TER 964 964
 SQ SEQUENCE 964 AA; 108287 MW; 90E72EDFAB1358B7 CRC64;

Query Match 23.9%; Score 1262.5; DB 6; Length 964;
 Best Local Similarity 33.4%; Pred. No. 1.9e-97;
 Matches 332; Conservative 169; Mismatches 349; Indels 143; Gaps 38;

QY 47 VGSKSSYPVWSEPEDLGCAALRQSSGTVEAAVEVDVSATITLOVLVDAPGNISCLWV 106
 DB 20 VQTGSSQPSV--SPEEL-----SPSIIHPAKS-ELIVSAGDEIRLFCITDPSVK--WT 67
 QY 107 FKHSLNCOFHPDLQNGVVMVLKMTQAGEYLLFIOSEATNYTILFTVSRN-TLL 165
 DB 68 FETLG-----QLSENTHAEIVKAEAMNTGNYCTNEGGLSSSIYVF---VRDPEKL 117
 QY 166 YTLRRYFRKMNQDALVCISVPPVIEVWVLCDSQGESCKESPAV-----VKKE 217
 DB 118 FLVDPPLYGK-EDNALVRCPLTDP-E-VTNYSLTGCCEGKPLKDLTFLVADPKAGITIRNV 175
 QY 218 EKVLHELFGTDIRC--CARNELGRE-CTRLFTID----LNQTTOTTLPO--LFLKVGEP 268
 DB 176 KEYHRL-----CLHCSANQGSKVSLSKFTLKVRAAIRAVPVAVSKASYLLRGE 229
 QY 269 WIRCKAVHVNHGFLTWELN---KALEGNYFEMSTYSTNRTMIRILPAFVSVARN 325
 DB 230 AVNCLIKDVSSSDVMWIRENSQTKAQVKNRSHQGDEN---FLRQERLTISSARVND 285
 QY 326 GYITCSKSKHPSQALVT-----IVGKGFINA-----TNSSEDYIDQYEEFCFSVR 373
 DB 286 GVPMCYANNITFGSANVTTLLEVVDKGFINIFPMNTTTFVNDGEDVDL-----IVEY 337
 QY 374 KAYPO-----IRCTWTSRKSFCPCQKGLDNGVYSIKFCNKH-----QPGYIF 418
 DB 338 EAYPKPEHROWIYNNRTATDKWEDYPKSE-----NESNIYVSELHUTRLKGTGGTYTF 392
 QY 419 HAENDDAQFTKMTLIRRPQVLA--EASASQASCFSDGYPLPSWTWKKCSKSPNCTE 476
 DB 393 LVSNADVNSVTNVYNTKPEILTHDRNLNMGMLQCVAAAGFPEPTIDWYFCPTGEQCSV 452
 QY 477 EI-TEGVNKRKANRVPGQWSSSTLMNSBAIKGLVKCAYNSLGTSCETILLNSPGPF 535
 DB 453 PVGPDVDQIQNSSVSPGKLVIHSSIDYSAFKINGTVECRAYNDVGKS--SAFFN-----F 506
 QY 536 PFTQD-----NISFATTGVCLLFIIVLTLLICHKKYKKQFRYESQLQWV-QVTGS 584
 DB 507 AFKEQIHAHTLFTPLLLIGFVIAAGMMCIIVMLT-----YKYLQPKMYEVQWKKVEEING- 561

QY 585 SDNEYFYVDFREYEDLKWEFPRENLEFGKVLGSGAGFKVMNATAYGISKTGYIOVAVK 644
 DB 562 --NNVYIIDPTQLPYDKHWEFPNRLSPFKTLGAGFGKVEATAYGLIKSDRAMTVAVK 619
 QY 645 MLKEKADSSERALMSLKMOTQLGSHENIVNLGACTLSGPIYLIPIEYCCYGDLLNLYR 704
 DB 620 MLKPSAHLTEREALMSLKVLSYLGNNHMINVLGACTTGGPTLVITEYCCYGDLLNLYR 679
 QY 705 SKREKPHRTWTE-----IPKEHNFYSYTFQSHPNSSM---PGSREVOIHPSDDQISGL 755
 DB 680 RKRDSPICQKQEDHAEALYKNLLHSKSSCSSTNEYMDMKPGVSIV-VPTKADKERSA 738
 QY 756 HGNSFHSEDEIEYENQKRLSEEDNLVLFEDLLCFAYQVAKGMEFLKFSVHRDLAAR 815
 DB 739 RIGSY-----IERDVTPIAMEDEL-ALDLELLSFSYQVAKGMAFLASKNCHIRDLAAR 792
 QY 816 NVLVTHGKVKIKCDFGLARDIMSDSNVYVRGNARLPVKWMAPESLFEGYITIKSDVMSYG 875
 DB 793 NILLTHGRITKICDFGLARDIKNDSNVYVKNARLPVKWMAPESEIFNCVYTFESDVMSYG 852
 QY 876 ILLWEIFSLGVNYPGIPVDANFYKLIQNGFKMDQPFYATEEYIIMQSCWAFDSRKPS 935
 DB 853 IFLWELFSLGSGPYPGMPVDSKPYKMKEGFRMLSPHAPAEYMDIMKTCWDADPLKRP 912
 QY 936 FPNLTSFLGCOLADAEAMYNQVGRVSEC-PH 967
 DB 913 FKQIVQLIEKQISESTNHIYSN----LANCSPH 941

RESULT 8
 Q8AXC6 PRELIMINARY; PRT; 984 AA.
 ID Q8AXC6;
 AC Q8AXC6;
 DT 01-MAR-2003 (TEMBLrel. 23, Created)
 DT 01-MAR-2003 (TEMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TEMBLrel. 23, Last annotation update)
 DE Kinase receptor C-kit.
 GN C-KIT.
 OS Fugu rubripes (Japanese pufferfish) (Takifugu rubripes).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
 OC Tetraodontidae; Tetraodontidae; Takifugu.
 OX NCBI_TaxID:31033;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Williams H., Brenner S., Venkatesh B.;
 RT "Characterization of the Platelet-derived Growth Factor Receptor Alpha
 and c-kit Genes in the Pufferfish Fugu rubripes.";
 RL DNA Seq. 13:263-270 (2002).
 DR EMBL; AF456419; AAN87555.1; -.
 KW Receptor.
 SQ SEQUENCE 984 AA; 110732 MW; 27D1EDC837CE8295 CRC64;

Query Match 23.9%; Score 1262; DB 13; Length 984;
 Best Local Similarity 38.2%; Pred. No. 2.2e-97;
 Matches 290; Conservative 114; Mismatches 255; Indels 100; Gaps 21;

QY 266 EPLWIRCKAVHVNHGFLTWEL---ENKALEGNYFEMSTYSTNRTMIRILPAFVSSVAR 322
 DB 220 ESYLTCNTNTVNGNTKLKWAPLGSQLPAKVDGS---SRILTENFTQARSATLHIAAVRI 276
 QY 323 NDTGYTCSKSKH---PSQSALVTIVGKGF---NATNSEDYIDQYEEFCFSVRKAYP 377
 DB 277 QDTGRVQCEAENEKGVSTQSVLDVEKGFMYSNPNVNG-TIQVRAGESLLLSVIEAYP 335
 QY 378 QIR-CTWTFSRKSPFCQKGLDN-----GYSIS---KFCNHK-HQPEYIFHAE 421
 DB 336 MPASASWSFW-----GRGLHNTSDHVTTRSHVTSSELKVLRLKMGEGVYTFQAS 388
 QY 422 NDDAQTKMTLIRNRPQVLAESA--SQASCFSDGYPLPSWTWKKCSKSPNCTEIT 479
 DB 389 NGDASVNHHTTIFVISKPIVSHGPDVGVRCVAGFPAPQITWYTCQFYARCQOVN 448

480 -----EGVNRKANRKYRGQWSSSTLNMBAIKGFLVKCCAYNSLGTSCETILLNSGP 534
449 ATQEQNVITVTLSPDFGKTEVESRVNISG--RFTLECUATVEGQATLFSISERT 506
535 F-----PFIQDNISFYATIGVCLLFIIVLLTLLCHYKQFRYESQLQMVQVGTSSDNE 588
507 ISHDLFSPLLGSVS-----AACILCLILVLF--YKYMOKPKYQIOWKVIE--GIHNN 557
589 YFYVDREYEDLWEPRENLEFGKVLGSGAFGKVMNATAYGSKTGVSIQVAVKMLKE 648
558 YVYIDPTQLPYDQHWEPFRKRLFGKTLGSGAFGKVEATAYGLANEDSMVTAVKMLKS 617
649 KADSSREALMSELKMTOLGSHENIVNLLGACTLSGPYLIIFYCCYCGDLNLYRSKRE 708
618 SAHSTEKALMSELKVLILYGNHINIVNLLGACTVGGTTLVITECCPGDLNLFRRKRE 677
709 KPHRTWTETFEKHNFSFY-----PTFQSHPNSSMPGSRVQTHP 747
678 SF-----ICPKLEBCHYRNTLMQREMGDSLNGVMTMRPSAAGKPSSSSSSEKRSILRE 732
748 DSDOISGLHGNSFHSEDEIEVENOKRLEEBEDNLVLPEDLLCFAYQVAKWMEFLPKSC 807
733 GSPY-----VEEDSESEMFDEDSLSLDT-EDLLSFYSYQVAKWMEFLTSKNC 777
808 VHRDLAARNVLVTHGKVKICDFGLARDIMSDSNVVRGNARLPVKWMAPELSLEGYTI 867
778 IHRDLAARNILLTQGRVAKICDFGLARDINTDSNVYVKGARLPVKWMSPEISFEVYTF 837
868 KSDVWSYGILLWEIPLSGVNPYPGIPVDANFYKLIQNGKMDQPFYATEEYIIMQSCWA 927
838 ESDVWSYGILLWEIPLSGVNPYPGIPVDANFYKLIQNGKMDQPFYATEEYIIMQSCWA 927
928 FDSKRPSFPNLTSFLGQCLADAEAMYNQVDRVSECP 966
898 ADPLNRPPFRKVERIEBQSLSDTTKHIYLNFSRVPVMP 936

RESULT 9
Q9TQ01 PRELIMINARY; PRT; 964 AA.
AC Q9TQ01;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Mast/stem cell growth factor receptor (Fragment).
OS KIT1*0201.
GN Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
SEQUENCE FROM N.A.
RP STRAIN=Yorkshire / Landrace cross; TISSUE=Leukocyte;
RC Marklund S.;
RL Thesis (1997), Department of Animal Breeding and Genetics,
RL Swedish University of Agricultural Sciences, Uppsala, SWEDEN.
RN [2]
SEQUENCE FROM N.A.
RP STRAIN=Yorkshire / Landrace cross; TISSUE=Leukocyte;
RC MEDLINE=98391767; PubMed=9724328;
RX Marklund S., Kijas J., Rodriguez-Martinez H., Roennstrand L., Puna K.,
RA Moller M., Edfors-Lilja I., Andersson L.;
RT "Molecular evolution of the dominant white phenotype in the domestic
pig";
RL Genome Res. 8:826-833(1998).
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-
CC PROTEIN KINASES.
DR EMBL: AJ223229; CAA11197.1; --
DR HSP: P11362; IFGK.
DR InterPro: IPR003599; Ig.
DR InterPro: IPR007110; Ig-like.

InterPro: IPR003006; Ig_MHC.
InterPro: IPR000719; Prot_kinase.
InterPro: IPR001824; RTKinaselli.
InterPro: IPR001245; Tyr_kinase.
Pfam: PF00047; Ig; 1.
DR Pfam: PF00069; pkinase; 1.
DR ProDom: PD000001; Prot_kinase; 2.
DR SMART: SM00409; IG_3_kinase; 2.
DR SMART: SM00219; Tyrc; 1.
DR PROSITE: PS00835; IG_Like; 3.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE: PS00240; RECEPTOR_TYR_KIN_III; 1.
KW ATP-binding; Glycoprotein; Immunoglobulin domain; Kinase;
KW Phosphorylation; Receptor; Tyrosine-protein kinase;
KW Tyrosine-protein kinase.
FT NON TER 964 964
SQ SEQUENCE 964 AA; 108315 MW; 996C346201358A8 CRC64;
Query Match 23.9%; Score 1261.5; DB 6; Length 964;
Best Local Similarity 33.4%; Pred. No. 2.3e-97;
Matches 332; Conservative 169; Mismatches 349; Indels 143; Gaps 38;
QY 47 VKSSSYPMVSESPEDIGCALRPOSSCTVYEAADVDSASITLQVLVDAPGNISCLWV 106
DB 20 VQTSSQPSV--SPEEL-----SPSIHPAKS-ELIVSAGDIRLFCCTDPGSVK--WT 67
QY 107 FKHSLLNCQPHFDLQNRGVVSMVILKMTETQAGYLLFIQSEATNYTLFTVSRN-TLL 165
DB 68 FETLG-----QLSENTHAEWIVEKAEAMNTGNYTCNEGGLSSIVF---VRDPEKL 117
QY 166 YTLRRPFRKVENODALVCISESVPEPIVEKVLCDSSQGESCKEESPAY-----VKKE 217
DB 118 FLVDPPLYGK-EDNDALVRCPDTPDPE-VTVNYSLTCGEGKPLKDLTFVADPKAGITIRNV 175
QY 218 EKVLEHFGTDIRC--CARNELGRE-CTRLFTID-----LNQTPQTLPO--LFLKVGEP 268
DB 176 KREYHRL-----CLHCSANQGGKSVLSKFTPLKVRRAAIRAVPVVAVSKASYLLREGSEF 229
QY 269 WIRKAVVNHGFLTWELN---KALEEGNYPEMSTYSTNRTWIRILFAVSSVARNDT 325
DB 230 AVMLIKDVSSVDSVMIRENSQTKAQVKNSSWHQGFN-----FLRQRLTISARVND 285
QY 326 GYTCSSSKHPSOSALVT---IVGKGFINA-----TNSSDEYEDIDQEEFCFSYRF 373
DB 286 GVFCYANNTFGSANVTTLLEVVDKFINIFPMNTTVFVNDGEDVDL-----IVEY 337
QY 374 KAYPO-----IRCTWTSRKSFPCEQKGLDNGYSIKFCNKH-----QPGEYIP 418
DB 338 EAYPKPEHRQWIMNRTATDKWEDYPKSE-----NESNIRYVSELHLTRLKGTGTYTF 392
QY 419 HAENDDAQFTKMTLNIIRKPOVLA--EASASQASCFSDGVPPLPSWTWKCKSDKSPNCTE 476
DB 393 LVSNADVNSSTVFNVTNTPKPEILLTHDRLMNGMLQCVAAAGFPEPTIDWYFCPTGTCQSCV 452
QY 477 EI-TEGVNRKANRKYRGQWSSSTLNMBAIKGFLVKCCAYNSLGTSCETILLNSGP 535
DB 453 PVGPDVQIQNSSVSPGKLVHSSIDYSAFKNGTVECRAYNDVGKS--SAFN-----F 506
QY 536 PFTOD-----NISFYATIGVCLLFIIVLLTLLCHYKQFRYESQLQMV-QVTGS 584
DB 507 AFKEQIHAHTLFTPLLIGFVIAAGMMCIIVMLT---YKYLQKPMVEVQWKKVEEING- 561
QY 585 SDNEYFVVDREYEDLWEPRENLEFGKVLGSGAFGKVMNATAYGSKTGVSIQVAVK 644
DB 562 --NNYVIDPTQLPYDQHWEPFRKRLFGKTLGSGAFGKVEATAYGLKISDAAMTAVK 619
QY 645 MLKEKADSSREALMSELKMTOLGSHENIVNLLGACTLSGPYLIIFYCCYCGDLNLYR 704
DB 620 MLKPSAHLTEREALMSELKVLVSLGNHINIVNLLGACTIGGTTLVITEYCCYCGDLNLYR 679
QY 705 SKREKPHRTWTE-----IFKEHNFSFYPTFQSHPNSSM---PGSREVQIHPDSQISGL 755

```

Db 680 RKDSDSFCSKQEDHAEALYKNLLHSSKSCSDSTNEYMDKPGSVY-VPTKADKERSA 738
QY 756 HGNFSHSEDEIEYENQKRLSEEDLNVLTFEDLLCFAYQVAKGMEFLFKSCVHRDLAAR 815
Db 739 RIGSY-----IERDVTPAIMEDDEL-ALDLEDLLSFYSQVAKGNWFLASKNCIHRDLAAR 792
QY 816 NVLVTGKVKIKCDFGLARDIMSDSNVVRGNARLPVKWMAPELSLFGIYTIKSDVMSYG 875
Db 793 NILTHGRIYKICDFGLARDIKNSNVVKGARLPVKWMAPELSLFGIYTIKSDVMSYG 852
QY 876 ILLWEIPLSGVNPYPGIPVDANFYKLIQNGFKMDQPFYATEEIIYIMQSWAFDSRKPS 935
Db 853 IFLWELFSLGSSYPGMPVDSKFKYMKIEGFRMLSPHAPVEMYDINKTCWDADPLKRP 912
QY 936 FPNLTSLGQLADAEABEAMQNDGRVSEC-PH 967
Db 913 FKQIVLIEKQISESTNHIYSN-----LANCSPH 941

RESULT 10
Q9TQ00
ID Q9TQ00 PRELIMINARY; PRT; 964 AA.
AC Q9TQ00;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Mast/stem cell growth factor receptor (Fragment).
GN KIT1*0202.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Yorkshire / Landrace cross; TISSUE=Leukocyte;
RA Marklund S.;
RL Thesis (1997), Department of Animal Breeding and Genetics,
RL Swedish University of Agricultural Sciences, Uppsala, SWEDEN.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Yorkshire / Landrace cross; TISSUE=Leukocyte;
RX MEDLINE=98391767; PubMed=9724328;
RA Marklund S., Kiljas J., Rodriguez-Martinez H., Roennstrand L., Funa K.,
RA Moller M., Edfors-Lilja I., Andersson L.;
RT "Molecular evolution of the dominant white phenotype in the domestic
RT pig.";
RL Genome Res. 8:826-833(1998).
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-
CC PROTEIN KINASES.
DR EMBL; AJ223230; CAAL1198.1; -.
DR HSSP; P11362; IFGK.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR001824; RTKinaseII.
DR InterPro; IPR001245; Tyr_kinase.
DR Pfam; PF00047; Ig; 1.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 2.
DR SMART; SM00409; IG; 3.
DR SMART; SM00219; TyrKc; 1.
DR PROSITE; PS00835; IG_LIKE; 3.
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE; PS00011; PROTEIN KINASE DOM; 1.
DR PROSITE; PS00109; PROTEIN KINASE TYR; 1.
DR PROSITE; PS00240; RECEPTOR_TYR_KIN_III; 1.
KW ATP-binding; Glycoprotein; Immunoglobulin domain; Kinase;
KW Phosphorylation; Receptor;
KW Tyrosine-protein kinase.
FT NON_TER , 964

```

```

SQ SEQUENCE 964 AA, 108287 MW, 1B21292A962E9191 CRC64;
Query Match 23.9%; Score 1261.5; DB 6; length 964;
Best Local Similarity 33.4%; Pred. No. 2.3e-97;
Matches 332; Conservative 169; Mismatches 349; Indels 143; Gaps 38;
QY 47 VGSSTSPWSSPELDGALRPQSSGTYYEAAAVEVDUSASITLQVLVDAPGNISCLWV 106
Db 20 VQTGSSQPSV--SPEEL-----SPPSIHAKS--ELIVSAGDEIRLFCFDPGSVK--WT 67
QY 107 FKHSLLNCQPHFLQNRGVVSVILKMTETQAGEYLLFIQSEATNYTILFTVSIRN--TLL 165
Db 68 FETLG-----QLSENTHAEWIEKAEAMTGNCTYCTNEGGLSSSIYVP--VRDEPKL 117
QY 166 YTLRRYFRMENQDALVCISESPPIVEMVLCDQSQGESCKEESPAP-----VKKE 217
Db 118 FLVDPPLYGK-EDNDALVRCPLDPE-VTNYSLTGCEGKPLPKDLTFVADPKAGITIRNV 175
QY 218 EKVLHELFGTDIRC--CARNELGRE-CTRLFTTD-----LNQTPQTLPO--LFLKVGEP 268
Db 176 KREYHRL-----CLHCSANOGGKSVLSKKFTLKVRAAIRAVPVVAVSKASYLLREGEEF 229
QY 269 WIRCKAVHNVHGFGLTWELN---KALEEGNYFEMSTYSTNRTMIRILFAFVSSVARNDT 325
Db 230 AVNCLIKDVSSVDSVMIRENSOTKAQVRKNSWHQGFN-----FLROEKUTISSARVND 285
QY 326 GYITCSTSSHPQSALVT---IVGKGFINA-----TNSSEDIYDQYEEFCFVRF 373
Db 286 GVFMVCYANNTFGSANVTTLLEVVDKGFNIFPMWNTTVFVNDGEDVDL-----IVEY 337
QY 374 KAVPQ-----IRCTWTSRSPFCQCKGLONGYSISKFNHKKH-----QPGYIF 418
Db 338 EAYPKPEHRQWIYMNRTATDKWEDYPKSE-----NESNIRYVSELHLTRLKGTGGTYTF 392
QY 419 HAENDDAQFTKMTLAIIRKPOVLA--EASASOASCFSDGCVPLSPSWTKKCDSPNCTE 476
Db 393 LVSNADVNSSVTFNYYNTKPEILTHDRLMNGLMQCVAAGFPETIDWYFCPTGEQRCSV 452
QY 477 EI-TEGVNRKANRKFVGQWSSSTLMSBAIKGFLVKCCAYNSLGTSCETILNSPGPF 535
Db 453 PVGPDVQIQNSSVSPFKLVHSSIDYSAFKENGTVCEKAYNDVGKS--SAFEN----F 506
QY 536 PFTQD-----NISFYATTGCVLLFIVLTLICHYKKQFRYESQLQWY-QVTGS 584
Db 507 AFKEQIHATLFTPLLLIGFVIAAGMMCIIVMLT---YKYLQKPMYEVQWKVVEEING- 561
QY 585 SDNEYFVVDPREYEDLKWEFPRENLEFGVLGSGAGFKVMNATAYGISKTGVSIQVAVK 644
Db 562 --NNYVIDPTQLPYDHKWEFPNRNLSFGTLGAGAFGKVEATAYGLIKSDAAMTAVK 619
QY 645 MLKEKADSSEREALMSLQMTQLGSHENIVNLIGACTLSGPIYLIFFEYCCYGDLNLYR 704
Db 620 MLKPSAHLTEREALMSLKVLSVGNHNVNLLGACTIGTPTLVITEYCCYGDLNLYR 679
QY 705 SKKEKPHRTWTE-----IFKEHNFSPYPTFQSHPNSSM---PGSRVQIHPDSDQISGL 755
Db 680 RKDSDSFCSKQEDHAEALYKNLLHSSKSCSDSTNEYMDKPGSVY-VPTKADKERSA 738
QY 756 HGNFSHSEDEIEYENQKRLSEEDLNVLTFEDLLCFAYQVAKGMEFLFKSCVHRDLAAR 815
Db 739 RIGSY-----IERDVTPAIMEDDEL-ALDLEDLLSFYSQVAKGNWFLASKNCIHRDLAAR 792
QY 816 NVLVTGKVKIKCDFGLARDIMSDSNVVRGNARLPVKWMAPELSLFGIYTIKSDVMSYG 875
Db 793 NILTHGRIYKICDFGLARDIKNSNVVKGARLPVKWMAPELSLFGIYTIKSDVMSYG 852
QY 876 ILLWEIPLSGVNPYPGIPVDANFYKLIQNGFKMDQPFYATEEIIYIMQSWAFDSRKPS 935
Db 853 IFLWELFSLGSSYPGMPVDSKFKYMKIEGFRMLSPHAPVEMYDINKTCWDADPLKRP 912
QY 936 FPNLTSLGQLADAEABEAMQNDGRVSEC-PH 967
Db 913 FKQIVLIEKQISESTNHIYSN-----LANCSPH 941

```



```

RESULT 11
Q8WN23 ID Q8WN23 PRELIMINARY; PRT; 979 AA.
AC Q8WN23;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE C-KIT.
OS
OC Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Pissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RA Zenke D., Yuzbasiyan-Gurkan V.;
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-
CC PROTEIN KINASES.
DR EMBL: AF448148; AAL40833.1;
DR InterPro; IPR003599; Ig.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003006; Ig MHC.
DR InterPro; IPR000719; Prot kinase.
DR InterPro; IPR001824; RTKinaselII.
DR InterPro; IPR001245; Tyr_kinase.
DR Pfam; PF00047; Ig; 2.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 2.
DR SMART; SM00409; IG; 3.
DR SMART; SM00219; Tyr_Kc; 1.
DR PROSITE; PS50835; IG_LIKE; 2.
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE; PS00011; PROTEIN KINASE DOM; 1.
DR PROSITE; PS00109; PROTEIN KINASE TYR; 1.
DR PROSITE; PS00240; RECEPTOR TYR_KIN III; 1.
KW ATP-binding; Glycoprotein; Immunoglobulin domain; Kinase;
KW Phosphorylation; Receptor; Transmembrane;
KW Tyrosine-protein kinase.
SQ SEQUENCE 979 AA; 109753 MW; 46C30D5DEB8E33D3 CRC64;

Query Match 23.78; Score 1250.5; DB 6; Length 979;
Best Local Similarity 33.68; Pred. No. 2e-96;
Matches 334; Conservative 161; Mismatches 345; Indels 153; Gaps 39;

QY 47 VKSSGYPVMSSEPDIGCALRPOSSGTVYEAADVEVDVSASITLQVLVDAPGNISCLWV 106
DB 22 VQTGSSQPSVSPGEPSP-LP-SIHPAKS-----ELIVSGDELRLSCTDPPGVK--WT 69
QY 107 FKH-SSLNCOPIHDLQNRGVSVMLKMTQTQAGEYLLFTIQSEATNYTIL-----FT 157
DB 70 FETLGLQNLNTH-----NEWITE-----KAEAG-----HTGNYTCTNRDGLSRISY 110
QY 158 VSRN-TLLYTLRRPFRKMNQDALVCISEPVEPIVWVLDQSQESKESPAV--- 213
DB 111 VFVRDPAKFLVDLPYLGK-EGNDTLVRCLTDP-E-VTNYSLRGCEGKPLKLTFFVADP 168
QY 214 -----VKKEKVLHLEFGTDIRC-CARNELGRE-CTRLEFTID-----LNQTPQTLTQL- 260
DB 169 KAGITIRNVKREYRL-----CLHSGADQKGRVLSKFTLKVRAAIRAVPVVSVSKTS 222
QY 261 -FLKVGEPILWIRKAVVHNGHGLTWELN-----KALEEGNYFEMSTYSTRNTRWIRLFA 315
DB 223 SLLKEGEAFVSMCFIKDVSSFVDSMWIKENSQQTNAQTQSNWHHGDFNERQEKLI--- 279
QY 316 FVSSVARNITGYTCCSSKHPQSALVT-----IVKGFINA-----TNSSEDEIDQ 363
DB 280 -ISSARNVDSGVFMCVANNITFGSANVTTLLEVVDVKGFINIFPMWSTTIFVNDGENVDL-- 336
QY 364 YEEFCFSVRFAKPYQ-IRCTWTSRKSFPCEQKGL---DNGYSISKFNCHK-----Q 412

```

```

337 -----IVEYBAYPKPEHQWIMYMRRTFTDKWEDYPKSDNESNI-RYVSELHLTRLKGN 389
413 PGEYIFHAENDDAQFTKMTLINIRKPKQVLAESASQA--SCFSDGYPLPSTWTKCSOK 470
390 GGTFTFQVNSDVNSVTFNVYVNTKPEILTHESLTNGMLQCWAGFPPEPAVDWYFCPGA 449
471 SPNCTEEI-TEGVNRKANRVFGQWSSSTLNMSEAIKFLVKCAVNSLGTSCETILL 529
450 EQRCSVPFGMDVQWNSLSPSGKLVVQSSIDYSAFKHNGIVECRAYNNVGRS--SAFP 507
530 NSPGPPFPFIQDN-----ISFYATIGVCLLFIIVLTLILCHYKKQKPFYESQ 575
508 N----FAFKGNSKEQLHPHTLFTPLLIGFVIAAGWMCIIIMLT-----YKYLQKPMYEQ 559
576 LQMV-QVTGSSDNEYFYVDREYEDLKWBPENLEFQKVLGSGAGPKVMNATAYGISK 634
560 WKVVEEING---NNYVYIDPTQLPYDHKEWPPNRNLSFGKTLGAGAFGKVVEATAYGLIK 616
635 TGVSIQVAVKMLKEKADSSEREALMSELKMTQLGSHENIVNLLGACTLSGPIYLIFEYC 694
617 SDAAMTVAVKMLKPSAHLTEREALMSELKVSUYLGNHNI VNLGACTVGGPTLVITEYC 676
695 CYGDLNLYLSRKREKFRHTWTE-----IFKEHNFSFYPTFQSHPNSSM---PGSREVOI 745
677 CYGDLNLFRLKRDSDFCCKQEDHGEVALYKNLLHKSCESSDSTNEYMDKPGVSYV-V 735
746 HPDSQISGLHNGSFHSEDEIEYENQKRLBEEDLNVLTFEDLLCFAYQVAKGMFLEPK 805
736 PTKADKRSARIGSY-----IERDVTPTAIMEDDEL-ALDLEDLLSFSYQVAKGMFLASK 789
806 SCVHRDLAARNVLVTHGVKVICDPLGLARDIMSDSYVVRGNARLPVTKWMAPESEFEGYI 865
790 NCIRDLAARNILLTHGRITKICDFGLARDIKNDSYVYVKGARLPVTKWMAPESEFNCYI 849
866 TIKSDVMSYGILLWEIPLSGVNPYPGPVDANFYKLIQNGFKMDQPPVATEIYIIMOSC 925
850 TPESDWSYGIPLWELFSLGSSPYGPNVDSKFTYMKIKEGFRMLSPERHAPAEYDIMKTC 909
926 WAFDSRKRPSFPNLTSLFGCLQADAEAMYNQV 958
910 WDADPLKRPFTKQIVQLIEKQISDSTNHIYSNL 942

RESULT 12
Q99662 ID Q99662 PRELIMINARY; PRT; 972 AA.
AC Q99662;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE KIT protein.
GN KIT.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92228497; PubMed=1373482;
RA Andre C., Martin E., Cornu F., Hu W.X., Wang X.P., Galibert F.;
RT "Genomic organization of the human c-kit gene: evolution of the
RT receptor tyrosine kinase subclass III.";
RL Oncogene 7:685-691(1992).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=97179223; PubMed=9027509;
RA Andre C., Hampe A., Lachaume P., Martin E., Wang X.P., Manus V.,
RA Hu W.X., Galibert F.;
RT "Sequence analysis of two genomic regions containing the KIT and the
RT FMS receptor tyrosine kinase genes.";
RL Genomics 39:216-226(1997).
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-

```

PROTEIN KINASES.

CC EMBL; U63834; AAC50969.1; --
DR HSSP; P11362; IFGK.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003598; Ig_c2.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR001824; RTK_inaseII.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00047; ig; 1.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 2.
DR SMART; SM00408; IGc2; 1.
DR SMART; SM00219; TyrKc; 1.
DR PROSITE; PS50835; IG LIKE; 1.
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE; PS00111; PROTEIN KINASE DOM; 1.
DR PROSITE; PS00109; PROTEIN KINASE_TYR; 1.
DR PROSITE; PS00240; RECEPTOR_TYR_KIN_III; 1.
KW ATP-binding; Glycoprotein; Immunoglobulin domain; Kinase;
KW phosphorylation; Receptor; Transferase; Transmembrane;
KW Tyrosine-protein kinase.
SQ SEQUENCE 972 AA; 109450 MW; D59DEFE9AF761PDA CRC64;

Query Match 23.6%; Score 1247; DB 4; Length 972;
Best Local Similarity 32.8%; Pred. No.4e-36;
Matches 323; Conservative 166; Mismatches 353; Indels 144; Gaps 34;

QY 47 V GKSSSSPMYSESPEDLGCALRPOSSGTVYEAADVDSASITLOVLVDAPGNISCLWV 106
DB 20 VQTGSSQPSVSPG-EPSPPSIHPQKSDLI-----VRVGDEIRLLCTDPGFVK--WT 67
QY 107 FKHSLLNCQPHFDLQNRGWSMWILKMTETQAGEYLLFIQSEATN---YTILFTVSIKNT 163
DB 68 FE-----ILDETENKQEWIT-EKAEATNTGKYCTNKHGLSNS 106
QY 164 L-----LYTLRRPYFRKMNQDALVCISVPPEIVBWVLCDSGESCKBE----- 209
DB 107 IYFVRDPAKLFLVDRSLYK-ENNDTLVRCLTDPE-VTNYSLKGCGKPLPKDLREIP 164
QY 210 ---SPAVVKEEKVLHFLFGTDIRCCARNELGRECTRLFTIDL---NOTQTLLPQ--L 260
DB 165 DPKAGIMIKSVKRAVHRLC---LHCSVDQEGKSVLSEKFIUKVAPAFKAVPVWSVKASY 221
QY 261 FLKVGPELWIRCKAVHNHGFGLWELEN--KALEEGNYFEMSTYSTNRTMIRILFAFV 317
DB 222 LLRGEFEFTVCTIKDYSSSVYSTWKRENSQTKLEKYNWHHGDFFNYERQAT----LTI 277
QY 318 SSVARNDTGYTCSSSKHPQSALVT---IVKGGFNA-----TNSEDEVIDOYE 365
DB 278 SSARVNSDGVFMCVANNFTFGSANVTTLTEVVDKGFNIFFPMINTVTFVNDGENVDL---- 333
QY 366 EFCFSVRFKAYPO-IRCTWTSTRSKSF--PCEQKGLDNGYSISKFCNHKH-----QDGE 415
DB 334 ---IVEYEAPFKPEHQOMIYNNRTFTDKWEDYPKSESNIRVYSELHLRLTKGTEGGT 389
QY 416 YIFHAENDDAQFTKMTFLNIRRKPOVLA--BASASQASCFSDGYPPLPSWTKKCSDKSPN 473
DB 390 YTFVLVNSDVMNAATAFNVVYNTKPEILLYDRLVNGMLQCVAGFPPEPIDWYFCGTQR 449
QY 474 CTBEITE-GVNNRKNRKNVQGVQSSSTLNMSSEAIKFLVKCCAYNSLGTGCTILLNSP 532
DB 450 CSASVLPVDVQTLNSSGPPFGKLIVQSSIDSSAFKHNGTVECKAYNDVGKT--SAYFN-- 505
QY 533 GPFFFIQD-----NISFYATIGVCLLFIIVVLTLLICHYKKQKPRYSOLQMV-QV 581
DB 506 --FAFKEQIHPTLFTPLLIGFVIWAGMCMCIIVMILT----YKYLQKPMYEQWKVVEEI 559
QY 582 TGSSDNEYFYVDFREYEDLKWEPRENLEFGKVLGSGAGFKVMNATAYGTSKTVSQV 641
DB 560 NG---NNVYIIDTQLPYDHWKEFFPNRLSPFKTLGAGAFKGVBATAYGLTKSDAAMTV 616
QY 642 AVKMLKEKADSSEREAALMSELKQMTQLGSHENIVNLLGACTLSGGPIYLIFCYCCYGDLLN 701

Matches	330;	Conservative	161;	Mismatches	346;	Indels	158;	Gaps	30;
Qy	47	VGKSSSYPMVSES	PEDLCALRPOSSGTYEAAAVEVDVSASITLQVLVDAPGNISCLWV	105					
Db	22	VQTGSSQPSVSPG	PSLP-SIHPAKS-----ELIVSGDELRLSCTDGFVK-WT	69					
Qy	107	FKH-SSNLQCPHDL	QNRGVVSMVLKMTETQAGEYLLFIOSEATNYTL-----FT	157					
Db	70	PETLGQNLNETH	---NEWITE-----KAEAG-----HTGNYTCTNRDGLSRSIY	110					
Qy	158	VSIRN-TLLYTLRR	PPYFRKMENQDALVCISBSVPPEPIVEMVLCDSQGSCKEESPAV---	213					
Db	111	VFVRDPAKLFVLD	PLYGK-EGNDTLVRCLPTDPE-VTNYSLRCEGKPLPKDLTFVADP	168					
Qy	214	-----VKEEKVL	HELFTDIRC-CARNELGRE-CTRLFTID-----INQTPQTTLPQL-	260					
Db	169	KAGITIRNVKREY	HRL-----CLHCSADQGRVTLSKKFTLVKRAIRAVPVVSVSKTS	222					
Qy	261	-FLKVGEP	LWTRCKAVHVNHGFLTWELN---KALEBGNFYEMSTYSTNRTMIRILFAF	316					
Db	223	SLLEGEAF	SVWCFIKDVSSPFDVSNWIKENSQTVAQTQSNWHHGGDFNFERQEKLI	278					
Qy	317	VSSVARNDTGY	YTCSSSHKPSQSALVT---IVGKGFINA-----TNSSEDEYIDQY	364					
Db	279	ISSARVND	SGVMFCYANNTFGSANVTTLLEVVDKGFINIPEMMSTTIFVNDGENVDL---	335					
Qy	365	EEFCFSVR	FKAYPO-IRCTWTFRSKSPCEOKGI---DNGYSISKPCNHKI	413					
Db	336	-----IVEYE	AYPPEHQOQTYMNRITDKWEDYPKSDNESNI-RYVSELTLRLKNGEG	389					
Qy	414	GEYIFHAEND	DAQTKFTLIRKRPOLAEASASQA--SCFSDGYPLPSPMTWKCKSDKS	471					
Db	390	GTYTFQV	SNDSVNSSVTFNVVNTKPELLTHESLTNGMLQCVAGFPPEPAVDWTFCEGAE	449					
Qy	472	PNCTEEI-TEG	VMNRKANRVKFGQWVSSTLNMGEAIKGFVLVCCAYNSLGTSGCETILN	530					
Db	450	QRCSVP	ICPMDVQMNSSLSPSGKLVVQSSIDYSAFKHNGTVCECRAYNNVGRS--SAFFN	507					
Qy	531	SPGPFPP	IQDN-----ISFYATIGVCLLPIVVLTLILICHKYKQFRRYSOL	576					
Db	508	-----FAFK	GNSSKEQIHPTHLPFLILGFIAGMMCIIMILT---YKLOKPMYEVQW	559					
Qy	577	QMV-QVTG	SSDNEYFYVDPRYEYDLKWEFPRENLEFKVLGSGAFGKWMATAYGISKT	635					
Db	560	KWVEE	ING--NNVYIDPTQLPYDHKWEFPNRNLSFGKTLGAGAFKVVEATYGLIKS	616					
Qy	636	GVSTQAV	KMLKEKADSSEREAALMSELKMMTQLGSHENIVNLGACTLSGPIYLIIFYCC	695					
Db	617	DAAMTV	AVKMLKPSAHLTEREALMSELKVLSYLGNHNNIVNLGACTVGGPTLVITEYCC	676					
Qy	696	YGLLNL	YLRSKREKFRHTWTWEIPEKH-NFSPYPTFQSHPNSSMPGSRREVQIHDPDSQISG	754					
Db	677	YGLLNL	FLRRKRDQSFICSKQE--DHGEVALYKMLLHSSKSSCCSDS-----TNEYMDM	726					
Qy	755	LHGNSF	HSDEIEYENOKLEE-----EEDLNLTFFEDLLCFAYQVAKGMFLE	803					
Db	727	KPGDSY	VVYPTKADKRRKRIGSYIERDVTPAIMEDDDELALDLELLSFSYQVAKGMAFLA	786					
Qy	804	FKCSV	HEDLAARNVLVTHGKVKVICDPLGARDIMSDSNVYVRGNARLPVKWMAPESLFEG	863					
Db	787	SKNCI	HRDLAARNILLTHGRITKICDPLGARDIKNDSYVYVKGARLPVKWMAPESEIFNC	846					
Qy	864	IYTIK	SDVMSYGILLWEIFSLGVNVPYPGIPVDANFYKLIQNGFKMDQPPFYATEIYIIMQ	923					
Db	847	VYTFES	DVMSYGIFLWELFSLGSSPYFGMPVDSKFYKWKIEGFRMLSPFHAPAEYDIMK	906					
Qy	924	SWAF	PSRKPSFNLTSFLGCLADAEAEAMYNQV	958					
Db	907	TCWD	ADPLKPTSKQIVQLIEKQISDSTNHIYSNL	941					

RESULT 14
Q98SU1

ID	Q98SU1	PRELIMINARY;	PRT;	977 AA.
AC	Q98SU1;			
DT	01-JUN-2001	(TREMBLrel. 17, Created)		
DT	01-JUN-2001	(TREMBLrel. 17, Last sequence update)		
DT	01-MAR-2003	(TREMBLrel. 23, Last annotation update)		
DE	Receptor tyrosine kinase Fms.			
OS	Danio nigrofasciatus.			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;			
OC	Cyprinidae; Danio.			
OX	NCBI_TaxID=144739;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=21359118; PubMed=11466528;			
RA	Patchy D.M., Johnson S.L.;			
RT	"Zebrafish hybrids suggest genetic mechanisms for pigment pattern			
RT	diversification in Danio.",;			
RL	Dev. Genes Evol. 211:319-328(2001).			
CC	-!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).			
CC	-!- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-			
CC	PROTEIN KINASES.			
DR	EMBL; AF324481; AAKI5303.1; --			
DR	HSSP; P11362; 1FGK.			
DR	InterPro; IPR007110; Ig-like.			
DR	InterPro; IPR003598; Ig C2.			
DR	InterPro; IPR003006; Ig MHC.			
DR	InterPro; IPR000719; Prot kinase.			
DR	InterPro; IPR001824; RtkKinaseIII.			
DR	InterPro; IPR001245; Tyr_pkinase.			
DR	Pfam; PF00047; Ig; 4.			
DR	Pfam; PF00069; pkinase; 1.			
DR	ProDom; PD000001; Prot kinase; 2.			
DR	SMART; SM00408; IGC2; 1.			
DR	SMART; SM00219; TyrKc; 1.			
DR	PROSITE; PS50835; IG_LIKE; 3.			
DR	PROSITE; PS00107; PROTEIN KINASE ATP; 1.			
DR	PROSITE; PS00117; PROTEIN KINASE DOM; 1.			
DR	PROSITE; PS00109; PROTEIN KINASE TYR; 1.			
DR	PROSITE; PS00240; RECEPTOR_TYR_KIN_III; 1.			
KW	ATP-binding; Glycoprotein; Immunoglobulin domain; Kinase;			
KW	Phosphorylation; Receptor; Transferrase; Transmembrane;			
KW	Thyrosine-protein kinase.			
FT	VARIANT 60 60 V -> L.			
FT	VARIANT 103 103 A -> S.			
FT	VARIANT 742 742 P -> S.			
FT	VARIANT 899 899 E -> D.			
SQ	SEQUENCE 977 AA; 109881 MW; CF54E129FAB10E3E CRC64;			
Query Match				
Best Local Similarity 23.3%; Score 1230.5; DB 13; Length 977;				
Matches 326; Conservative 174; Mismatches 367; Indels 127; Gaps 34				
Qy	71	SSGTVYEAARVVDVSASITQLVLDAPGNISCL-WYFKGSSLCQPHFDLQNRGVSMV	129	
Db	27	NSGAL---AGTDVLDGSGPLQVCEGDGPFVTPRVAKH-----KRYISKEVGKIRSF	78	
Qy	130	ILKWTETQAGEY-LLFTQSEATNYTLTFSIRNT-LLY-----TLRPFYRKMNQDAL	182	
Db	79	VERATVDFTGYKCVYINGSDSNLASSVHFVDRSVLFLVSPSSLR--YVRK-EGEDLL	135	
Qy	183	V-CISESVPEP-IVEWVLCDSQGESCKEESPVAVKKEKVL---HELFGTDIRCCARNE	236	
Db	136	LPCL---LTPDATDFTFRMDNGSAAPYGNNAFDPRKGVLIIRNVHGFNADYICSAIG	192	
Qy	237	LGRECTRLFTDLNQTPQTTLPOLFK-----VGEPLWIRCAVHNHGFGLTWLEN	289	
Db	193	GAEKVSXIFSINVIQLRFF-PYVYLKRNEYKVLGGERLQISCTTNPNFYVNTWTHSS	251	
Qy	290	KALEEGNYFMWSTYSTNRTMIRILFAFVSSVARNDCTYTCSSKHPQSALVT---IVG	346	
Db	252	KRLPKPE--SKSTMEDRLAIESILT-ISSVQLSDTGNTICTGNEAGANSSITQLLAVD	308	
Qy	347	KGPI-----NATNSEDSEYIDQEEFCFSVRKAYPOIRC-TW---TFSRKSPFCEQ	394	

```
Db 309 EPIRLSPKLSKTHRLGSLIEVSEGEDVDGLVLEAYPPLTSHTWETPTSHNASLP--- 365
Qy 395 KGLDNGYSISFCNKH-----HOPGEYIFAENDDAQFTKMTLNRKRPQ 440
Db 366 -----ENRFYNHNDREYALLFLKRLNFEIGQVTLNVKMSKASITFDIKMYTPEV 417
Qy 441 VLAE-ASASQASCFSDGYPLPSWTWKCDKSPNCTEITE-----GVNKRKANRKFVQGW 495
Db 418 ARVKWENVTTLSRSGYGPAPSIWVQCTGRTTTPENTDLPQTQOTQVFEKESFGAV 477
Qy 496 VSSSTLNMSEAIKGLVKCCAYNSLGTSCETILLNSPGPPPIQDNISFYATIG--VCLL 553
Db 478 GVESVLTVPGN--RRMTVVCAFNLVGQSDTFSMD-----VSDQIFTSACGSGTVAMV 529
Qy 554 FIVVLTLICHYKQFQYESQLQWVOVTVGSSDNEYFVYDPEYEDLKWEPRENLEFG 613
Db 530 VLGLLILFIWYIKQKRYEIRKWLIEAT--NGNNTYTFIDPTQLPYNKWEFPRDKLKG 587
Qy 614 KVLGSAGFKVMNATAYGISTKGVSIQVAVKMLKEKADSSEREALMSLKMVTOLGSHEN 673
Db 588 KTLGAGAKGVVEATAYGLKEDNITRVAVKMLKASAHDPDEREALMSLKLILSHLGQKN 647
Qy 674 IVNLGACTLSPYILFEYCCYGLLVLRKREKPHRTWTETIPEKHNFSPYPTFQSHP 733
Db 648 IVNLGACTHGGPVLVITEYCCYGLLVLRKREKPHRTWTETIPEKHNFSPYPTFQSHP 733
Qy 734 NSSMP-----GSEVQIHPSDOIQLGHLGNSFHSEDEIYENKRLSE-----BEDLNV 782
Db 693 NPPEVTYKYNVSTERMVFRSDSGSSTCSHYLDMRPVTSRPTNSALDPSDCQEDSWP 752
Qy 783 LTFEDLLCFAYQVAKGMFLFEKSCVHRDLAARNVLVTHGKVKICDFGLARDIMSDSNY 842
Db 753 LDMDDLAFSSQVAGLOFLAAKNCIHRDVAARNVLLTNSRVAKICDFGLARDIMSDSNY 812
Qy 843 VVRGNARLPVKWAPESLFEYITKSDVWSYGIILLWEIFSLGVNYPGIPVDANFYKLI 902
Db 813 VVKGARLPVKWAPESLFEYITKSDVWSYGIILLWEIFSLGVNYPGIPVDANFYKLI 902
Qy 903 QNGFKMDOPFATEIYIIMQSCWAFDSKRPSPNLTSLFCQLADAEAM-----YQNV 958
Db 873 KCGYQMSRPFDFAPPENYIMKVCWNLAEARPTFSKIQMORMLGETSEQDQTOEYKNI 932
Qy 959 DGRVSECPHTYQNRPPFGREMDLGLLSPOAQVED 992
Db 933 PSE-AEVQPLESCDPVKEHDESFETSCEPEED 965

RESULT 15
O77589
ID O77589 PRELIMINARY; PRT; 945 AA.
AC O77589;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Tyrosine kinase receptor homolog (Fragment).
GN KIT.
OS Equus caballus (Horse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX NCBI_TaxID=9796;
RN [1].
RP SEQUENCE FROM N.A.
RC TISSUE=Skin;
RA Lear T.B., Adams M.H., McDowell K.J., Ponce de Leon F.A.,
RA Sullivan N.D., Cogle L., Ferguson E., Chambers T.M., Bailey E.;
RT "Chromosomal location of genes for ESR, ETS2, GOT2, KIT, MX1 and PGR
RT in the horse, Equus caballus.";
RL Proc. Xth Amer. Coll. Dom. Anim. Cytogenet. Gene Mapp. 0:0-0(1997).
RN [2].
RP SEQUENCE FROM N.A.
RC TISSUE=Skin;
RA Rieder S., Checa-Cortes M.L., Joerg H., Stranzinger G.;
```

```
RT "An Equine sequence homologous to the c-KIT-proto-oncogene mapped to
RL chromosome 3q13-21.";
RT Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE CSP-1/PDGF RECEPTOR FAMILY OF TYROSINE-
CC PROTEIN KINASES.
DR EMBL; AF055037; AAC23611.1; -.
DR HSSP; P11362; IFGK.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003006; Ig MHC.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR001824; RYKaseIII.
DR InterPro; IPR001245; Tyr_kinase.
DR Pfam; PF00047; Ig; 1.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 2.
DR SMART; SM00409; IG; 3.
DR SMART; SM00219; Tyrc; 1.
DR PROSITE; PS50835; IG_LIKE; 3.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00101; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE; PS00240; RECEPTOR_TYR_KIN_III; 1.
KW ATP-binding; Glycoprotein; Immunoglobulin domain; Kinase;
KW Phosphorylation; Receptor; Transferase; Transmembrane;
KW Tyrosine-protein kinase.
FT NON_TER 945
SQ SEQUENCE 945 AA; 106143 MW; 7DB424A89DF5C340 CRC64;

Query Match 23.3%; Score 1229; DB 6; Length 945;
Best Local Similarity 32.5%; Pred. No. 1.3e-94;
Matches 323; Conservative 166; Mismatches 368; Indels 136; Gaps 31;

Qy 47 VKSSSYPMWSESPDLGALRQSSGTVYEAAYVEVDVVSASITQLVLDAPGNISCLWV 106
Db 20 VQTGSQSPV--SPGEL-----SPPSIHPAKS-ELIVSGDEIRLLCADGFVK--WT 67
Qy 107 FK-HSSLNCQPHFDIQNRGVSMVILKMTETQAGEYLLFIQSEATNYTLFTVIRNTLL 165
Db 68 FETWQLSENTHKE-----WVTEKAEATNGSYTCTNEGGLSSIYVF---VRDPAK 116
Qy 166 YTLRRPYPRKMNQALVCISVPEPIVWVLCDSQGSCKEESPAV-----VKKE 217
Db 117 LFLFDPISLYGKESDITLVRCLTDPE-VTNYSIMACEGSKLPKDLTFVADPKAGITIRNV 175
Qy 218 EKVHLEHFGDIRCCARNELGRECTRLFTID---LNQTPQTLPO--LFLKVGGEPLWIR 271
Db 176 KREYHRLC---LRCSADKDGKSVLSNKKFTLVRAIRAVPVVSVSKASYLLREGGEFSVT 232
Qy 272 CKAVHNVHFGTLWELENKALEEGNVFEMSTYSTNRTMIRILPAFVSSVARNDTGYTCS 331
Db 233 CLIKDVSSVSDVMWIRENSRTKE-QVKSSSSHQGFNFVROERLTISPARVNDSGVFMCY 291
Qy 332 SKHPSQSALVT---IVGKGFINA-----TNSSEYIEDQYEEFCFSVRFKAYPQ- 378
Db 292 ANNTFGSANVTTLVLEVDKGCINVPFMNTTVFVNDGENVDL-----IVEYESYKPK 343
Qy 379 -----IRCTWTFPSRKPFPCEQKGLDNGYSISKCNKH-----QGEYIFHAENDD 424
Db 344 EHQQWIYMNRTSTDKEWDPKSE-----NESNIRYVSELHLTRLKGTGEGTYTFLVNSD 398
Qy 425 AQFTKMFNTLNRKPKQVLA--EASASQASCFSDGYPLPSWTWKCDKSPNCTEEL-TEG 481
Db 399 VDSSVTFNVTNKPEILTRDLNMGMLQCVAAAGEPEPTIDWYFGCTQRCSIPGVPD 458
Qy 482 VNRKANRKFVQGWVSSSTLNMSEAIKGLVKCCAYNSLGTSCETILLNSPGPPPIQDN 541
Db 459 VKIQNSSVPFGKLVVQSSIDYSAFKHGTVECRAYNDVGKSSAPF-----N 505
Qy 542 ISFYATIGVCLLF-----IVVLLILCHYKQFQYESQLQWV-OVTSSDN 587
Db 506 PAFKEQIHPHTLFTPLLIGSVVAAGMCMCVIMVLTLYKYLQKPMYEQWKVEEING---N 562
```

```

Qy 588 EYFVDFREYEDLKWEPPRENLEFGKVLGSGAFKGMNATAYGISKTGVSIQVAVKMLK 647
Db 563 NYVYIDPTQLPYDHKWEPPRNLSFGKTLGAGAFKGVVEATAYGLIKSDAAMTAVKMLK 622
Qy 648 EKADSSREALMSLKMTQLGSHENIVNLGACTLSGPIYLIIPYCCYGDLLNVLRSKR 707
Db 623 PSAHLTREALMSLKVLISYLGHNHNVNLGACTVGGTTLVITEYCCYGDLLNVLRSKR 682
Qy 708 EKFRHTWTE-----IPKEHNFSFYPTFQSHPNSSM---PGSREVOIHPDSQISGLHGN 758
Db 683 DSFTCSKQEDHAEALYKNLLHSHKSSCNDSTNEYMDMKPGVSYV-VPTKTKRRAARIG 741
Qy 759 SFHSEDIENQKRLBEEEDLNVLTPEDLCPAYQVAKGMFLEFKSCVHRDLAARNVL 818
Db 742 SY-----IERDVAPSIMEDDEL-ALDLEDLSFSYQVAKGMWAFKASKNCIHRDLAARNIL 795
Qy 819 VTHCKVVKICDFGLARDIMSDSNVVRGNARLPVKWMAPELSFEGIYTIKSDVWSYGILL 878
Db 796 LTHGRITKICDFGLARDIKNSDNVVRGNVRLPVKWMAPESIFENCYVTFESDVWSYGIFL 855
Qy 879 WEIFSLGWNYPGIPVDANFYKLIQNGFKMDQPFYATEEIIYIMQSCWAPDSRKRPSFPN 938
Db 856 WELFSLGSSYPGMPVDSKFKYKMIKEGFRMLSPHAPAEYDIMTKCWDADPLKRPTFKQ 915
Qy 939 LTSFLGQQLADAEEMAYQNVDRVSECPHTYQN 971
Db 916 IVOLIEKQISDSTNQIYSN-----LANCSPROEN 944

```

Search completed: August 26, 2003, 07:27:52
Job time : 87.5421 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2003, 07:13:34 ; Search time 62.5315 Seconds
(without alignments)
2520.578 Million cell updates/sec

Title: US-09-919-408A-4

Perfect score: 5274

Sequence: 1 MPALARDACTVPLLVFSAM.....PFSREMDLGLSPQAQVEDS 993

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5274	100.0	993	16 AAR67816	Flk2 receptor prot
2	5274	100.0	993	16 AAR67536	Human flk-2. Homo
3	5274	100.0	993	17 AAR97419	Murine foetal live
4	5274	100.0	993	18 AAW19873	Human flk-2 recept
5	5274	100.0	993	20 AAY08617	Human flk-2 protei
6	5274	100.0	993	23 ABG70916	Human receptor pro
7	5274	100.0	993	23 AAR25819	Human receptor pro
8	5271	99.9	14 AAR37503	Human flk-2. Homo	
9	5266	99.8	14 AAR44995	Human flk-2 recept	

10	5266	99.8	993	16 AAR75961	Human STX-1. Homo
11	5265	99.8	993	16 AAR81869	Human Flk2/flt3 ty
12	5262	99.8	1167	14 AAR31376	Human flk-2. Homo
13	5159	97.8	983	19 AAW63588	Human receptor typ
14	5157.5	97.8	986	19 AAW63587	Human receptor typ
15	5157.5	97.8	986	19 AAW63589	Human receptor typ
16	5153.5	97.7	994	19 AAW63586	Human receptor typ
17	4533.5	86.0	1000	16 AAR81868	Flk2/flt3 tyrosine
18	4429.5	84.0	992	16 AAR37502	Murine flk-2. Mus
19	4429.5	84.0	992	16 AAR67815	Flk2 receptor prot
20	4429.5	84.0	992	16 AAR67535	Mouse flk-2. Mus
21	4429.5	84.0	992	17 AAR97418	Human foetal liver
22	4429.5	84.0	992	18 AAW19874	Murine flk-2 recep
23	4429.5	84.0	992	20 AAY08616	Murine flk-2 prote
24	4429.5	84.0	992	23 ABG70915	Mouse receptor pro
25	4429.5	84.0	992	23 AAR25818	Murine receptor pr
26	4421.5	83.8	992	14 AAR44994	Murine flk-2 recep
27	4421.5	83.8	992	14 AAR31375	Murine flk-2. Mus
28	4404.5	83.5	992	13 AAR28038	Murine flk-2. Mus
29	3323	63.0	665	19 AAW63585	Human receptor typ
30	2747	52.1	749	19 AAW78002	Protein pMON32390.
31	1640	31.1	481	15 AAR47579	Soluble Flk-2. Mu
32	1578	29.9	307	23 ABB81198	Human FLT3 protein
33	1286	24.4	977	21 AAY51322	Bovine c-kit bk-1
34	1266	24.0	975	22 AAE07144	Murine Kit/stem ce
35	1266	24.0	975	22 AAE07148	Mutant murine Kit/
36	1251	23.7	976	22 AAE07145	Human Kit/stem cel
37	1251	23.7	976	22 AAE07149	Mutant human Kit/s
38	1251	23.7	976	22 AAU00375	Human stem cell gr
39	1251	23.7	976	24 ABP72761	Human c-kit. Homo
40	1251	23.7	976	24 ABR39481	Human c-kit wild-t
41	1251	23.7	976	24 ABP72341	Human c-kit. Homo
42	1251	23.7	976	24 ABP72342	Human c-kit. Homo
43	1251	23.7	976	24 ABP57411	Human c-kit protei
44	1251	23.7	976	24 ABP57418	Human c-kit protei
45	1251	23.7	976	24 ABP96025	Human c-kit protei

ALIGNMENTS

RESULT 1

AAAR67816	AAAR67816 standard; Protein; 993 AA.
ID	AAAR67816 standard; Protein; 993 AA.
XX	AC AAR67816;
XX	AC
XX	AC
DT	25-MAR-2003 (updated)
DT	18-AUG-1995 (first entry)
XX	Flk2 receptor protein-tyrosine-kinase.
DE	Human Flk2; receptor protein-tyrosine-kinase; primitive
KW	hematopoietic cell; fetal liver kinase; diagnostic ligand
KW	isolation; bone marrow disease therapy.
XX	Homo sapiens.
OS	Homo sapiens.
XX	
PH	Key
FT	Peptide
FT	/note= "signal peptide"
FT	Protein
FT	/note= "mature protein"
FT	Domain
FT	/note= "extracellular receptor domain"
FT	Domain
FT	/note= "transmembrane region"
FT	Domain
FT	/note= "intracellular catalytic domain"
XX	
XX	W09500554-A2.
PN	
XX	05-JAN-1995.
PD	

XX 17-JUN-1994; 94WO-US06944.
XX 18-JUN-1993; 93US-0080244.
XX 21-JUN-1993; 93US-0081508.
XX 23-NOV-1993; 93US-0157490.
XX (UYPR-) UNIV PRINCETON.
XX Lemischka IR;
XX WPI; 1995-052014/07.
XX N-PSDB; AAQ81013.
XX Ligand for receptor protein tyrosine kinase - useful for the
XX stimulation of primitive haematopoietic stem cells causing
XX proliferation and/or differentiation
XX Disclosure; Fig 1b; 13lpp; English.
XX The sequence corresponds to a human Flk2 (fetal liver kinase)
XX receptor protein-tyrosine-kinase, which is expressed in primitive
XX hematopoietic cells but not in mature hematopoietic cells. The
XX protein is useful in isolation of receptor ligands, which have
XX applications in diagnosis of bone marrow disorders and in
XX stimulating proliferation and/or differentiation of primitive
XX hematopoietic stem cells.
XX (Updated on 25-MAR-2003 to correct PN field.)
XX Sequence 993 AA;
XX
XX Query Match 100.0%; Score 5274; DB 16; Length 993;
XX Best Local Similarity 100.0%; Pred. No. 0;
XX Matches 993; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MPALADAGTVPVLLVFSAMIFGTITNQDLPVTKVLIHKNNDSSVCKSSSPMVSESP 60
DB 1 MPALADAGTVPVLLVFSAMIFGTITNQDLPVTKVLIHKNNDSSVCKSSSPMVSESP 60
QY 61 EDLGCALRPOSSGTVEAAAVEVDVSASITLQVLVDAPGNISCLWVFKHSLNCQPHFDL 120
DB 61 EDLGCALRPOSSGTVEAAAVEVDVSASITLQVLVDAPGNISCLWVFKHSLNCQPHFDL 120
QY 121 QNRGVSMVILKMTETQAGEYLLFIQSEATNYTLFTVSRNTLLYTLRRPYFRKMNQD 180
DB 121 QNRGVSMVILKMTETQAGEYLLFIQSEATNYTLFTVSRNTLLYTLRRPYFRKMNQD 180
QY 181 ALVCISESVPEPIVWVLCDSQGESCKEESPAVVKKEKVLHELFGTDIRCCARNEIGRE 240
DB 181 ALVCISESVPEPIVWVLCDSQGESCKEESPAVVKKEKVLHELFGTDIRCCARNEIGRE 240
QY 241 CTRLFTDLNQTPTTLPOLFLKVGEPFLWIRCKAVHNVHGFGLTWELNKALEBGNFYFEM 300
DB 241 CTRLFTDLNQTPTTLPOLFLKVGEPFLWIRCKAVHNVHGFGLTWELNKALEBGNFYFEM 300
QY 301 STYSTNRMTIRILFAFVSSVARNDGTYYTSCSSKHPQSQSALVTIVGKGFINATNSSDYE 360
DB 301 STYSTNRMTIRILFAFVSSVARNDGTYYTSCSSKHPQSQSALVTIVGKGFINATNSSDYE 360
QY 361 IDQYEEFCFSVRKAYPQIRCTWTFSRKSPCEQKGLDNGYSISKFCNHKHQPGYEYIFHA 420
DB 361 IDQYEEFCFSVRKAYPQIRCTWTFSRKSPCEQKGLDNGYSISKFCNHKHQPGYEYIFHA 420
QY 421 ENDDAQFTKMTLIRKPKQVLAASASQASCFSDGYPSPWTWKCKSDKSPNCTEITE 480
DB 421 ENDDAQFTKMTLIRKPKQVLAASASQASCFSDGYPSPWTWKCKSDKSPNCTEITE 480
QY 481 GVNMRKANRVFGQWVSSSTILNMSIAIKGFLVKCCAYNSLGTSCETILLNSPGFPFIQD 540
DB 481 GVNMRKANRVFGQWVSSSTILNMSIAIKGFLVKCCAYNSLGTSCETILLNSPGFPFIQD 540
QY 541 NISFYATIGVCLLFIVVLTLLICHYKQFRYESQLQMVQVTGSSDNEYFYVDPREYED 600

DB 541 NISFYATIGVCLLFIVVLTLLICHYKQFRYESQLQMVQVTGSSDNEYFYVDPREYED 600
QY 601 LKWEPPRENLEFGKVLGSGAFGKVMATAYGISTGTGVSIOVAVKMLKEKADSSEREALMS 660
DB 601 LKWEPPRENLEFGKVLGSGAFGKVMATAYGISTGTGVSIOVAVKMLKEKADSSEREALMS 660
QY 661 ELKMMTOLGSHENIVNLLGACTLSGPIYLIIFYCCYGGDLLNLYRSKREKPHRTWTEIFKE 720
DB 661 ELKMMTOLGSHENIVNLLGACTLSGPIYLIIFYCCYGGDLLNLYRSKREKPHRTWTEIFKE 720
QY 721 HNFSFYPTQSHPNSSMPGSRREVQIHPDSDQISGLHGNFSHSEDEIEYENQKLEEEEDL 780
DB 721 HNFSFYPTQSHPNSSMPGSRREVQIHPDSDQISGLHGNFSHSEDEIEYENQKLEEEEDL 780
QY 781 NVLTFTEDLLCFAYQVAKGMFELEFKSCVHRDLAARNVLTGHGVKVICDFGLARDINMSD 840
DB 781 NVLTFTEDLLCFAYQVAKGMFELEFKSCVHRDLAARNVLTGHGVKVICDFGLARDINMSD 840
QY 841 NYVVRGNARLPVKWMAPESLFEGIYTIKSDVMSYGILLWEIFSLGVNYPYGPVPDANFYK 900
DB 841 NYVVRGNARLPVKWMAPESLFEGIYTIKSDVMSYGILLWEIFSLGVNYPYGPVPDANFYK 900
QY 901 LIQNGFKMQPPFYATEEIIYIMOSWAFDSRKRPPNLTSLFLGCOLADAEAEAMYQNVDG 960
DB 901 LIQNGFKMQPPFYATEEIIYIMOSWAFDSRKRPPNLTSLFLGCOLADAEAEAMYQNVDG 960
QY 961 RVSECPHTYQNRPPFRSREMDLGLLSPQAQVEDS 993
DB 961 RVSECPHTYQNRPPFRSREMDLGLLSPQAQVEDS 993
RESULT 2
AAR67536
ID AAR67536 standard; Protein; 993 AA.
XX AAR67536;
XX 25-MAR-2003 (updated)
DT 04-JUL-1995 (first entry)
XX Human flk-2.
XX Fetal liver kinase-2; flk-2; protein tyrosine-kinase receptor;
KW hematopoiesis; stem cell.
XX Homo sapiens.
XX Key Location/Qualifiers
FT Peptide 1..27 /label= Sig_peptide
FT Domain 28..543 /label= Extracellular_receptor_domain
FT Region 544..563 /label= Transmembrane_region
FT Domain 564..993 /label= Intracellular_catalytic_domain
XX US5367057-A.
XX 22-NOV-1994.
XX 30-APR-1993; 93US-0055269.
XX 02-APR-1991; 91US-0679666.
XX 28-JUN-1991; 91US-0728913.
XX 15-NOV-1991; 91US-0793065.
XX 24-DEC-1991; 91US-0813593.
XX 26-JUN-1992; 92US-0906397.
XX 12-NOV-1992; 92US-0975049.
XX 19-NOV-1992; 92US-0977451.
XX 30-APR-1993; 93US-0055269.
XX (UYPR-) UNIV PRINCETON.
XX PA

XX PI Lemischka IR;
 XX WPI; 1995-005894/01.
 DR N-PSDB; AAQ79069.
 XX
 PT Murine flk-2 receptor protein tyrosine kinase - used to stimulate
 PT proliferation and/or stimulation of primitive mammalian
 PT haematopoietic stem cells in vitro or in vivo.
 XX
 PS Disclosure; Fig. 2A-1F; 69pp; English.
 XX
 CC cDNAs encoding receptor protein tyrosine-kinases, mouse fetal liver
 CC kinase-2 (flk-2), human flk-2 and mouse flk-1 are given in AAQ79068-70,
 CC respectively, and the deduced amino acid sequences in AAQ67535-37,
 CC respectively.
 CC (updated on 25-MAR-2003 to correct PF field.)
 XX
 SQ Sequence 993 AA;
 Query Match 100.0%; Score 5274; DB 16; Length 993;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 993; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MPALARDAGTVPLLVFSAIFGTITNODLPVVKVLIHKNNDSSVCKSSSYPMVSESP 60
 DB 1 MPALARDAGTVPLLVFSAIFGTITNODLPVVKVLIHKNNDSSVCKSSSYPMVSESP 60
 QY 61 EDLGCALRPOSSGTVYEAADVEVDVASITLQVLVDAPGNISCLWVFKHSLNCPHF 120
 DB 61 EDLGCALRPOSSGTVYEAADVEVDVASITLQVLVDAPGNISCLWVFKHSLNCPHF 120
 QY 121 QNRGVSMVILKMTQAGEYLLFIQSEATNTYTLFTVSRINTLLYTLRPPYFRKMNQD 180
 DB 121 QNRGVSMVILKMTQAGEYLLFIQSEATNTYTLFTVSRINTLLYTLRPPYFRKMNQD 180
 QY 181 ALVCISSEVPEPIVWVLCDSQGESCKEESPAVVKKEKVLHFLFGTDIRCCARNE 240
 DB 181 ALVCISSEVPEPIVWVLCDSQGESCKEESPAVVKKEKVLHFLFGTDIRCCARNE 240
 QY 241 CTRFTIDLNQTPOTLPQFLKVGEPWIRKCAVHNHGLTWELNKALEGNFEM 300
 DB 241 CTRFTIDLNQTPOTLPQFLKVGEPWIRKCAVHNHGLTWELNKALEGNFEM 300
 QY 301 STYSTNRTMIRILPAFVSSVARNDTGYTCSSSKHPSQSALVTIVGKGFINATNSS 360
 DB 301 STYSTNRTMIRILPAFVSSVARNDTGYTCSSSKHPSQSALVTIVGKGFINATNSS 360
 QY 361 IDQYEEFCFVRKAYPOIRCTWTFRSKPPCEQKGLDNGYSISKPCNKHQPGGEYIF 420
 DB 361 IDQYEEFCFVRKAYPOIRCTWTFRSKPPCEQKGLDNGYSISKPCNKHQPGGEYIF 420
 QY 421 ENDDAQFTKFTLNIRKPKQVLAESASQASCFSDGYPLPSWTWKKCDKSPNCTEITE 480
 DB 421 ENDDAQFTKFTLNIRKPKQVLAESASQASCFSDGYPLPSWTWKKCDKSPNCTEITE 480
 QY 481 GWNRKANRKFVGOWSSSTLWMEAIKFLVKCAVNSLGTSCETILLNSPGPPFTQD 540
 DB 481 GWNRKANRKFVGOWSSSTLWMEAIKFLVKCAVNSLGTSCETILLNSPGPPFTQD 540
 QY 541 NISFYATIGVCLLFIIVLTLIICHYKQFPYBSQOLQWQVGTGSDNBYFYVDFREYD 600
 DB 541 NISFYATIGVCLLFIIVLTLIICHYKQFPYBSQOLQWQVGTGSDNBYFYVDFREYD 600
 QY 601 LKWEFFRENLEFGVLGSGAFKWNATAYGISTGYSIOVAVKMLKEKADSSEREALMS 660
 DB 601 LKWEFFRENLEFGVLGSGAFKWNATAYGISTGYSIOVAVKMLKEKADSSEREALMS 660
 QY 661 ELKQMTOLGSHENIVLLGACTLGGPIYLIIEYCCYGDLLNLRKREKPHRTWTEIPKE 720
 DB 661 ELKQMTOLGSHENIVLLGACTLGGPIYLIIEYCCYGDLLNLRKREKPHRTWTEIPKE 720
 QY 721 HNFSFYPTFQSHPNSSMPGSRREVQIHPSDQISGLHGNSFHSDEIEYENQKRLSEEDL 780

DB 721 HNFSFYPTFQSHPNSSMPGSRREVQIHPSDQISGLHGNSFHSDEIEYENQKRLSEEDL 780
 QY 781 NVLTPEDLCLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVKIKCDFGLARDIMSDS 840
 DB 781 NVLTPEDLCLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVKIKCDFGLARDIMSDS 840
 QY 841 NYVVRGNARLPVVKWMAPESEFLFEGITTIKSDVWSYGILLWEIFSLGVNYPGIPVDANFYK 900
 DB 841 NYVVRGNARLPVVKWMAPESEFLFEGITTIKSDVWSYGILLWEIFSLGVNYPGIPVDANFYK 900
 QY 901 LIQNGFKMDQPPYATEIEIIMQSCWAFDSKRKPSFENLTSFLGCQLADAEAMYNQVNDG 960
 DB 901 LIQNGFKMDQPPYATEIEIIMQSCWAFDSKRKPSFENLTSFLGCQLADAEAMYNQVNDG 960
 QY 961 RVSECPHTYQNRPPFSREMDLGLLSPQAQVEDS 993
 DB 961 RVSECPHTYQNRPPFSREMDLGLLSPQAQVEDS 993
 RESULT 3
 AAR97419
 ID AAR97419 standard; Protein; 993 AA.
 XX
 AC AAR97419;
 DT 25-MAR-2003 (updated)
 DT 11-DEC-1996 (first entry)
 XX
 XX Murine foetal liver kinase 2.
 XX
 KW Murine; foetal liver kinase 2; flk-2; protein tyrosine kinase;
 KW monoclonal; antibody; extracellular domain; receptor assay;
 KW haematopoietic stem cell; ligand; stimulation; proliferation;
 KW differentiation; treatment; anaemia; bone marrow damage;
 KW cancer chemotherapy; radiation.
 XX
 OS Mus musculus.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..27
 FT Peptide /label= sig_peptide
 FT Peptide 28..993
 FT Peptide /label= mat_peptide
 FT Domain 28..544
 FT Domain /label= extracellular_domain
 FT Domain 545..564
 FT Domain /label= transmembrane_domain
 FT Domain 565..993
 FT Domain /label= intracellular_domain
 XX
 PN US5548065-A.
 XX
 PD 20-AUG-1996.
 XX
 PF 31-OCT-1994; 94US-0252517.
 XX
 PR 19-NOV-1992; 92US-0977451.
 PR 02-APR-1991; 91US-0679666.
 PR 28-JUN-1991; 91US-0728913.
 PR 15-NOV-1991; 91US-0793085.
 PR 24-DEC-1991; 91US-0813593.
 PR 26-JUN-1992; 92US-0906397.
 PR 12-NOV-1992; 92US-0975049.
 PR 30-APR-1993; 93US-0055269.
 PR 31-OCT-1994; 94US-0252517.
 XX
 PA (UYPR-) UNIV PRINCETON.
 XX
 PI Lemischka IR;
 XX
 DR WPI; 1996-392678/39.
 DR N-PSDB; AAT38734..

XX Anti-fetal liver kinase 2 (flk-2) antibodies - useful in assays,
PT for isolating haematopoietic stem cells expressing receptor and for
PT obtaining ligands
XX
PS Claim 1; Columns 39-48; 50pp; English.
XX
CC The present sequence is murine foetal liver kinase 2 (flk-2),
CC a protein tyrosine kinase. Isolated antibodies, pref. monoclonal,
CC raised against the extracellular portion of flk-2 can be used to
CC assay for flk receptors on the surface of primitive haematopoietic
CC stem cells, and to isolate positive cells. The antibodies can also
CC be used as, or to obtain ligands, which stimulate the proliferation
CC and/or differentiation of stem cells. The ligands can be used, e.g.
CC for treating anaemia, or bone marrow damage resulting from cancer
CC chemotherapy, or radiation.
CC (Updated on 25-MAR-2003 to correct PF field.)
XX
SQ Sequence 993 AA;

Query Match 100.0%; Score 5274; DB 17; Length 993;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 993; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MPALARDAGTVPLLVFSGAMIFGTTITNODLPVVKVNLNHNKNDSSVGKSSSYPMVSESP 60
Db 1 MPALARDAGTVPLLVFSGAMIFGTTITNODLPVVKVNLNHNKNDSSVGKSSSYPMVSESP 60
Qy 61 EDLGCALRQSSGTYEAAAEVDVSASITLQVLVDAPGNISCLVWFKHSSLNCPHFDL 120
Db 61 EDLGCALRQSSGTYEAAAEVDVSASITLQVLVDAPGNISCLVWFKHSSLNCPHFDL 120
Qy 121 QNRGVSVMLKMTQAGEYLLFTQSEATNTYILFTVSRNTLLYLRPRYFRKWNQD 180
Db 121 QNRGVSVMLKMTQAGEYLLFTQSEATNTYILFTVSRNTLLYLRPRYFRKWNQD 180
Qy 181 ALVCISESVPEPIVWVLCDSGESCKEESPAVKKEKVLHELFGTDIRCCARNELGRE 240
Db 181 ALVCISESVPEPIVWVLCDSGESCKEESPAVKKEKVLHELFGTDIRCCARNELGRE 240
Qy 241 CTRFTIDNCTPOTTLQFLKVGCEPLWIRCKAVVNHGFLTWELNKALEEGNYEM 300
Db 241 CTRFTIDNCTPOTTLQFLKVGCEPLWIRCKAVVNHGFLTWELNKALEEGNYEM 300
Qy 301 STYSTNRTMIRILFAFVSVAARNDTGYTTCSSSKHPSQSALVTIVGKGFINATNSSEDE 360
Db 301 STYSTNRTMIRILFAFVSVAARNDTGYTTCSSSKHPSQSALVTIVGKGFINATNSSEDE 360
Qy 361 IDQVEEPCFVRFKAYPOIRCTWTSPRSKSPCEQKGLDNGYSISKFCNHKHOPGEYIFHA 420
Db 361 IDQVEEPCFVRFKAYPOIRCTWTSPRSKSPCEQKGLDNGYSISKFCNHKHOPGEYIFHA 420
Qy 421 ENDDAQFTMTFLNIRRRKPOVLAEASQASCFSDGYPLPSWTWKCDKSPNCTEEITE 480
Db 421 ENDDAQFTMTFLNIRRRKPOVLAEASQASCFSDGYPLPSWTWKCDKSPNCTEEITE 480
Qy 481 GWNRRKANRKFQGVQWSSSTLNMSBAIKGFLVKCCAYNSLGTSCETILNPSGPPFFIQD 540
Db 481 GWNRRKANRKFQGVQWSSSTLNMSBAIKGFLVKCCAYNSLGTSCETILNPSGPPFFIQD 540
Qy 541 NISFYATIGVCLLFIWLTLLCHIKYKQFRIESQLOMVQVTGSSDNEYFYVDFREYED 600
Db 541 NISFYATIGVCLLFIWLTLLCHIKYKQFRIESQLOMVQVTGSSDNEYFYVDFREYED 600
Qy 601 LKWEFPRENLEPGKVLGSGAFQKVMNATYISKTGVSIOQAVKMLKEKADSSEREALMS 660
Db 601 LKWEFPRENLEPGKVLGSGAFQKVMNATYISKTGVSIOQAVKMLKEKADSSEREALMS 660
Qy 661 ELKQMTQIGSHENIYNLIGACTLSGPYILFYCCVCCGDLNLYLRKREKFTWTETPEKE 720
Db 661 ELKQMTQIGSHENIYNLIGACTLSGPYILFYCCVCCGDLNLYLRKREKFTWTETPEKE 720
Qy 721 HNFSFYPTFQSHPNSSMPGSRVQIHDPDSQIISGLHGNSFHSDEIEYENQKRLBEEDL 780

Db 721 HNFSFYPTFQSHPNSSMPGSRVQIHDPDSQIISGLHGNSFHSDEIEYENQKRLBEEDL 780
Qy 781 NVLTFEDLLCFAYQVAKGMEFELEFKSCVHRDLAARNVLVTHGKVVKI CDFGLARDIMSDS 840
Db 781 NVLTFEDLLCFAYQVAKGMEFELEFKSCVHRDLAARNVLVTHGKVVKI CDFGLARDIMSDS 840
Qy 841 NYVVRGNARLPVKWMAPELSELEFEGYTIKSDVMSYGILLWEIFSLGVNYPGIPVDANFYK 900
Db 841 NYVVRGNARLPVKWMAPELSELEFEGYTIKSDVMSYGILLWEIFSLGVNYPGIPVDANFYK 900
Qy 901 LIQNGFKMDQPFYATEEYIIIMQSCWAFDSRKRSPFNLTSLFLGCOLADAEAMYQNVDG 960
Db 901 LIQNGFKMDQPFYATEEYIIIMQSCWAFDSRKRSPFNLTSLFLGCOLADAEAMYQNVDG 960
Qy 961 RVSECPHTYQNRPRPFSREMDLGLLSPQAQVEDS 993
Db 961 RVSECPHTYQNRPRPFSREMDLGLLSPQAQVEDS 993

RESULT 4
AAW19873
ID AAW19873 standard; Protein; 993 AA.
XX
AC AAW19873;
DT 25-MAR-2003 (updated)
DT 19-AUG-1997 (first entry)
DE Human flk-2 receptor.
XX
KW Human; fetal liver kinase 2; flk2; receptor protein tyrosine kinase;
KW ptk; liver; spleen; thymus; adult; brain; bone marrow; primitive portion;
KW haematopoietic hierarchy; extracellular domain; soluble form; ligand;
KW proliferation; differentiation; mammalian; haematopoietic stem cell;
KW macrocytic anaemia; aplastic anaemia; cancer; radiation; mouse.
OS Homo sapiens.
FH Key Location/Qualifiers
FT Peptide 1..27 /note= "Signal peptide"
FT Protein 28..993 /note= "Mature flk-2"
FT US5621090-A.
PN 15-APR-1997.
PD 26-JUN-1992; 92US-0906397.
PF 26-JUN-1992; 92US-0906397.
PR 02-APR-1991; 91US-0679666.
PR 28-JUN-1991; 91US-0728913.
PR 15-NOV-1991; 91US-0793065.
PR 24-DEC-1991; 91US-0813593.
PA (UYPR-) UNIV PRINCETON.
XX Lemischka IR;
PI WPI; 1997-235228/21.
DR N-PSDB; AAT72117.
XX Protein containing the extracellular domain of human flk-2 - used
PT for identification of primitive haematopoietic cell proliferation
PT and differentiation stimulatory ligands, e.g. for treating anaemia
XX Claim 1; Fig 1B; 55pp; English.
PS This sequence represents human fetal liver kinase 2 (flk2). flk-2 is
CC a receptor protein tyrosine kinase (ptk) and is important in transducing
CC putative self-renewal signals from the environment. flk-2 is expressed

CC in fetal liver, spleen, fetal thymus, adult brain and adult bone marrow,
CC and it is thought that flk-2 is expressed in the entire primitive portion
CC of the haematopoietic hierarchy. The invention concerns a recombinant
CC nucleic acid, preferably mRNA, which encodes a protein containing only
CC the extracellular domain of human flk-2 and lacking the flk-2 intra-
CC cellular catalytic domain. The resultant protein represents a soluble
CC form of flk-2 which is used to isolate specific ligands for flk-2. These
CC ligands can be used to stimulate proliferation and/or differentiation of
CC mammalian haematopoietic stem cells, in vivo or in vitro, e.g. for
CC treatment of macrocytic or aplastic anaemia or bone marrow damage caused
CC by cancer treatment or radiation.
CC (Updated on 25-MAR-2003 to correct PP field.)
XX
SQ Sequence 993 AA;

Query Match 100.0%; Score 5274; DB 18; Length 993;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 993; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPALARDAGTVPLLVVFSAMIFGTTNQDLPVVKVLIHKNQNDSSVGKSSYPMVSESP 60
DB 1 MPALARDAGTVPLLVVFSAMIFGTTNQDLPVVKVLIHKNQNDSSVGKSSYPMVSESP 60

QY 61 EDLGCALRPOSSGVYEAAYVEVDVSAITLQVLVDAPGNISCLWFKHSLNCQPHFDL 120
DB 61 EDLGCALRPOSSGVYEAAYVEVDVSAITLQVLVDAPGNISCLWFKHSLNCQPHFDL 120

QY 121 QNRGVSWVLKMTQAGEVLLFIQSEATNTYTLFTVSIIRNTLLYTLRRYFRKMNQD 180
DB 121 QNRGVSWVLKMTQAGEVLLFIQSEATNTYTLFTVSIIRNTLLYTLRRYFRKMNQD 180

QY 181 ALVCISESVPEPIVEWVLCDSQGSCKEESPAVVKKEKVLHFLFGTDIRCCARNELGRE 240
DB 181 ALVCISESVPEPIVEWVLCDSQGSCKEESPAVVKKEKVLHFLFGTDIRCCARNELGRE 240

QY 241 CTRFLTIDLNQTPOTTLPOLFLKVGEPWIRCKAVHNVHGFGLTWELNKALEGNYPFEM 300
DB 241 CTRFLTIDLNQTPOTTLPOLFLKVGEPWIRCKAVHNVHGFGLTWELNKALEGNYPFEM 300

QY 301 STYSTNRTMIRILFAFVSSVARNDGYVTCSSSKHPSQALVTIVGKGFINATNSSDYE 360
DB 301 STYSTNRTMIRILFAFVSSVARNDGYVTCSSSKHPSQALVTIVGKGFINATNSSDYE 360

QY 361 IDQVEEFCFVRKAYPOIRCTWTFSRKSPFCEQKGLDNGYSISKFCNHKHPQGEYIFHA 420
DB 361 IDQVEEFCFVRKAYPOIRCTWTFSRKSPFCEQKGLDNGYSISKFCNHKHPQGEYIFHA 420

QY 421 ENDDAQFTKMTLNIIRKPOVLAEASASQCSFSDGYPLPSWTWKKCSKSPNCTEITE 480
DB 421 ENDDAQFTKMTLNIIRKPOVLAEASASQCSFSDGYPLPSWTWKKCSKSPNCTEITE 480

QY 481 GVMNRKANRKFQGVWSSSTLNMSAIGFLVKCCAYNSLGTSCETILLNSPGPPFIQD 540
DB 481 GVMNRKANRKFQGVWSSSTLNMSAIGFLVKCCAYNSLGTSCETILLNSPGPPFIQD 540

QY 541 NISFVATTGVCLLFTWLTLICHYKQFRYESQLQVQVGTSSDNEYFYVDREYD 600
DB 541 NISFVATTGVCLLFTWLTLICHYKQFRYESQLQVQVGTSSDNEYFYVDREYD 600

QY 601 LKWEPPRENLEFGKVGSAFGKVMNATAYISKTGVSIVQAVKMLKEKADSSEREAALS 660
DB 601 LKWEPPRENLEFGKVGSAFGKVMNATAYISKTGVSIVQAVKMLKEKADSSEREAALS 660

QY 661 ELKMTQLGSHENIVNLGACTLSGPILYIFRYCCYGDLLNLYLRKREKFTHTWTEIKE 720
DB 661 ELKMTQLGSHENIVNLGACTLSGPILYIFRYCCYGDLLNLYLRKREKFTHTWTEIKE 720

QY 721 HNFSPYPTFQSHPNSSMPSGRVQIHPDSQDQISGLHNSFHSDEIEYENQKLEEEEDL 780
DB 721 HNFSPYPTFQSHPNSSMPSGRVQIHPDSQDQISGLHNSFHSDEIEYENQKLEEEEDL 780

QY 781 NVLTPEDLILCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVVKICDFGLARDIMSDS 840

DB 781 NVLTPEDLILCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVVKICDFGLARDIMSDS 840
QY 841 NYVVRGNARLPVKWMAPESLFEGIYTIKSDVWSYGILLWEIFSLGVNPPYGPVPDANFYK 900
DB 841 NYVVRGNARLPVKWMAPESLFEGIYTIKSDVWSYGILLWEIFSLGVNPPYGPVPDANFYK 900
QY 901 LIQNGFMDQOPFYATBEIYIIMQSWAFDSKRKPSFNLTSFLGCQLADAEAMYQNVGD 960
DB 901 LIQNGFMDQOPFYATBEIYIIMQSWAFDSKRKPSFNLTSFLGCQLADAEAMYQNVGD 960
QY 961 RVSECPHTYQNRPPFSREMDLGLLSPOAQVEDS 993
DB 961 RVSECPHTYQNRPPFSREMDLGLLSPOAQVEDS 993

RESULT 5
AAAY08617
ID AAY08617 standard; Protein; 993 AA.
XX
AC AAY08617;
XX DT 05-AUG-1999 (first entry)
XX XX Human flk-2 protein.
XX DE Human flk-2; flk-1; cell isolation; fetal liver kinase; receptor;
XX KW monoclonal; polyclonal; antibody; tyrosine kinase.
XX OS Homo sapiens.
XX XX US5912133-A.
XX PD 15-JUN-1999.
XX XX 10-FEB-1998; 98US-0021324.
XX PR 19-NOV-1992; 92US-0977451.
XX PR 02-APR-1991; 91US-0679666.
XX PR 28-JUN-1991; 91US-0728913.
XX PR 15-NOV-1991; 91US-0793065.
XX PR 24-DEC-1991; 91US-0813593.
XX PR 26-JUN-1992; 92US-0906397.
XX PR 12-NOV-1992; 92US-0975049.
XX PR 30-APR-1993; 93US-0055269.
XX PR 31-OCT-1994; 94US-0252498.
XX PR 15-FEB-1996; 96US-0601891.
XX (UYPR-) UNIV PRINCETON.
XX Lemischka IR;
XX WPI; 1999-357194/30.
XX N-PSDB; AAX77515.
XX Isolating hematopoietic cells expressing fetal liver kinase 1
PT receptors
XX
PS Disclosure; Fig 1b; 59pp; English.
XX
CC This invention describes a novel method of isolating cells expressing
CC fetal liver kinase 1 (flk-1) receptors on their surface and comprises
CC binding the cells to a polyclonal or monoclonal antibody specific to the
CC the flk-1 receptor and isolating the cells that have bound to the
CC antibody. The method can be used to isolate hematopoietic stem cells in
CC any mammal but preferably a rat, mouse, rabbit or human. The proteins of
CC the invention belong to the receptor protein family. This sequence
CC represents the human flk-2 protein which is used in the method of the
CC invention.
XX
SQ Sequence 993 AA;

Query Match 100.0%; Score 5274; DB 20; Length 993;
Best Local Similarity 100.0%; Pred. No. 0;

Matches	993;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
Qy	1	MPALARDAGTVLLVVFSA	MPGTTITN	QDLPV	KVCLIN	HNKNDSSV	GKSSYP	PMVSESP	60
Db	1	MPALARDAGTVLLVVFSA	MPGTTITN	QDLPV	KVCLIN	HNKNDSSV	GKSSYP	PMVSESP	60
Qy	61	EDLGCALRQSSGTVYEA	AAVEVDV	SASIT	LOVLV	DAPGNISCL	WVFKHSS	LNCPHF	120
Db	61	EDLGCALRQSSGTVYEA	AAVEVDV	SASIT	LOVLV	DAPGNISCL	WVFKHSS	LNCPHF	120
Qy	121	QNRGVVSVILKMTQAGE	YLLFTQ	SEATNT	ILFTV	SRINTL	YTLRRP	FRKME	180
Db	121	QNRGVVSVILKMTQAGE	YLLFTQ	SEATNT	ILFTV	SRINTL	YTLRRP	FRKME	180
Qy	181	ALVCISESVPEIVFWL	CDGCKES	PAVVK	KEKVL	HELFG	TDIRC	CARNEL	240
Db	181	ALVCISESVPEIVFWL	CDGCKES	PAVVK	KEKVL	HELFG	TDIRC	CARNEL	240
Qy	241	CTRLEITIDNQTPOT	TLQFLK	VGCEPL	WIRCKA	VHVHGF	GLTWEL	ENKALEE	300
Db	241	CTRLEITIDNQTPOT	TLQFLK	VGCEPL	WIRCKA	VHVHGF	GLTWEL	ENKALEE	300
Qy	301	STYSTNRTMIRILPA	FVSSVAR	NDGYT	CSKSP	QSALVT	IVGKF	FINATN	360
Db	301	STYSTNRTMIRILPA	FVSSVAR	NDGYT	CSKSP	QSALVT	IVGKF	FINATN	360
Qy	361	IDQYEEFCFVRFKAY	POIRCT	TWTSR	KSFPC	EQKGL	NGYSI	KFCNKH	420
Db	361	IDQYEEFCFVRFKAY	POIRCT	TWTSR	KSFPC	EQKGL	NGYSI	KFCNKH	420
Qy	421	ENDDAQFTMTFLN	RRKPOV	LAESA	QSCF	SDGYPL	PSWTW	KKCDK	480
Db	421	ENDDAQFTMTFLN	RRKPOV	LAESA	QSCF	SDGYPL	PSWTW	KKCDK	480
Qy	481	GWNKRNKRVFGQW	SVSSTL	NMSAI	KGLVK	CCAYNS	LGTS	CTILLNS	540
Db	481	GWNKRNKRVFGQW	SVSSTL	NMSAI	KGLVK	CCAYNS	LGTS	CTILLNS	540
Qy	541	NISFYATIGVCLL	FIWVLT	LLI	CHYK	KQFRY	ESQ	LMQVQ	600
Db	541	NISFYATIGVCLL	FIWVLT	LLI	CHYK	KQFRY	ESQ	LMQVQ	600
Qy	601	LKWEPFRENLEFG	KVLGSG	AGFKV	NNATAY	GISKT	GVSIQ	VAVMK	660
Db	601	LKWEPFRENLEFG	KVLGSG	AGFKV	NNATAY	GISKT	GVSIQ	VAVMK	660
Qy	661	ELKMMTQLGSHEN	IYNL	GACTL	SGPIV	LI	REYCC	VDLLN	720
Db	661	ELKMMTQLGSHEN	IYNL	GACTL	SGPIV	LI	REYCC	VDLLN	720
Qy	721	HNFSFYPTFQSH	PNMSPG	SREVQ	IHPDS	DQIS	GLHGS	FHSEDE	780
Db	721	HNFSFYPTFQSH	PNMSPG	SREVQ	IHPDS	DQIS	GLHGS	FHSEDE	780
Qy	781	NVLTFEDLLCFAY	QVAKW	MEFL	EFKSC	VHRDLA	ARNVL	THGKV	840
Db	781	NVLTFEDLLCFAY	QVAKW	MEFL	EFKSC	VHRDLA	ARNVL	THGKV	840
Qy	841	NVVRGNARLPVK	WAPSL	PEGI	YTIK	SDVWS	GILLW	EISL	900
Db	841	NVVRGNARLPVK	WAPSL	PEGI	YTIK	SDVWS	GILLW	EISL	900
Qy	901	LIQNGFKMDQFP	YATEE	IYIM	QSCW	AFDSR	KRPFP	NLNTS	960
Db	901	LIQNGFKMDQFP	YATEE	IYIM	QSCW	AFDSR	KRPFP	NLNTS	960
Qy	961	RVSECPHYQNR	RPPSRE	MDLGL	LSLP	QAQV	EDS	993	
Db	961	RVSECPHYQNR	RPPSRE	MDLGL	LSLP	QAQV	EDS	993	

ID	ABG70916 standard; Protein; 993 AA.
XX	
AC	ABG70916;
XX	
DT	09-DEC-2002 (first entry)
XX	
DE	Human receptor protein tyrosine kinase, FLK-2.
XX	
KW	Human; FLK-2; foetal liver kinase; antianaemic; enzyme;
KW	ophthalmological; receptor protein tyrosine kinase; aplastic anaemia;
KW	primitive haematopoietic cell; stem cell; macrocytic anaemia;
XX	bone marrow damage; cancer chemotherapy.
OS	Homo sapiens.
XX	
FH	Key
FT	Peptide
FT	Protein
FT	Protein
XX	
PN	US2002119545-A1.
XX	
PD	29-AUG-2002.
XX	
PF	01-JUN-2001; 2001US-0872136.
XX	
PR	19-NOV-1992; 92US-0977451.
PR	30-APR-1993; 93US-0055269.
PR	31-OCT-1994; 94US-0252498.
PR	15-FEB-1996; 96US-0601891.
PR	10-FEB-1998; 98US-0021324.
PR	10-DEC-1998; 98US-0208786.
PR	02-APR-1991; 91US-0679666.
PR	28-JUN-1991; 91US-0728913.
PR	15-NOV-1991; 91US-0733065.
PR	24-DEC-1991; 91US-0813593.
XX	
PA	(LEMI/) LEMISCHKA I R.
XX	
PI	Lemischa IR;
XX	
DR	WPI; 2002-731356/79.
DR	N-PSDB; ABS55045.
XX	
PT	New protein tyrosine kinase expressed in primitive hematopoietic cells
PT	(HC) and not expressed in mature HC, and ligands for the protein,
PT	useful for stimulating proliferation of primitive hematopoietic stem
PT	cells -
XX	
PS	Claim 37; Fig 1b; 64pp; English.
XX	
CC	The invention relates to a receptor protein tyrosine kinase (pTK)
CC	expressed in primitive haematopoietic cells and not expressed in mature
CC	haematopoietic cells, and named FLK-1 and -2 (foetal liver kinase).
CC	Also included are the nucleic acids encoding the FLK proteins, FLK
CC	expression vectors, a ligand that binds to human or murine FLK-2 or FLK-1
CC	and stimulates the proliferation and/or differentiation of the primitive
CC	haematopoietic cells and a murine cell line 2018 having American type
CC	Culture Collection (ATCC) accession number ATCC CRL 10907. The ligands
CC	are useful for stimulating the proliferation and/or differentiation of
CC	primitive mammalian haematopoietic stem cells. The receptor pTK molecules
CC	are useful for stimulating the self-renewal of the totipotent
CC	haematopoietic stem cell and to stimulate the development of all cells of
CC	the haematopoietic system both in vitro and in vivo. The ligands for the
CC	receptors act as haematopoietic growth factors. The ligands are useful in
CC	treating humans whose primitive stem cells do not sufficiently undergo
CC	self-renewal e.g. macrocytic and aplastic anaemia and bone marrow damage
CC	resulting from cancer chemotherapy and radiation. The present
CC	sequence represents human FLK-2.
XX	
XX	Sequence 993 AA;

Query Match		100.0%;	Score 5274;	DB 23;	Length 993;
Best Local Similarity		100.0%;	Pred. No. 0;		
Matches 993;		Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	MPALARDAGTVPLLVFVSAMIFGTTNQDLFVIRKVLINHKNDSSVKGSSSYPMVSESP	60		
Db	1	MPALARDAGTVPLLVFVSAMIFGTTNQDLFVIRKVLINHKNDSSVKGSSSYPMVSESP	60		
Qy	61	EDLGCALRPOSSGTVYAAAVEVDVSAITLOVLVDAPGNISCLWVFKHSLNCQPHDL	120		
Db	61	EDLGCALRPOSSGTVYAAAVEVDVSAITLOVLVDAPGNISCLWVFKHSLNCQPHDL	120		
Qy	121	QNRGVSMWILKMTETQAGEVLLFIQSEATNTILFTVSIIRNTLLYTLRRPYFRMENQD	180		
Db	121	QNRGVSMWILKMTETQAGEVLLFIQSEATNTILFTVSIIRNTLLYTLRRPYFRMENQD	180		
Qy	181	ALVCISSEVPPIVEWVLCDSQGSCKEESPAVVKKEKVLHELFGTDIRCCARNELGRE	240		
Db	181	ALVCISSEVPPIVEWVLCDSQGSCKEESPAVVKKEKVLHELFGTDIRCCARNELGRE	240		
Qy	241	CTRLFTIDLNQTPQTTLPQLFKVGEPLWIRCKAVHVNHGFLTWELNKALEBGNYPFM	300		
Db	241	CTRLFTIDLNQTPQTTLPQLFKVGEPLWIRCKAVHVNHGFLTWELNKALEBGNYPFM	300		
Qy	301	STYSTNRTMIRILPAFVSSVARNDGYTTCSSKHPQSALVTIVGKFINATNSSEDEYE	360		
Db	301	STYSTNRTMIRILPAFVSSVARNDGYTTCSSKHPQSALVTIVGKFINATNSSEDEYE	360		
Qy	361	IDQYEEFCFSVRKAYPOIRCTWTFRSKSPFCEOKGLDNGYSISKFCNHKHPQGEYIFHA	420		
Db	361	IDQYEEFCFSVRKAYPOIRCTWTFRSKSPFCEOKGLDNGYSISKFCNHKHPQGEYIFHA	420		
Qy	421	ENDDAQTKMFTLNIRRKPOVLAEASASQSCFSDGYPLPSWTWKCKSDKSPNCTEITE	480		
Db	421	ENDDAQTKMFTLNIRRKPOVLAEASASQSCFSDGYPLPSWTWKCKSDKSPNCTEITE	480		
Qy	481	GVMNRKANRVFGQWVSSSTLWNGEAIKGFVLKCCAYNSLGTSCBTILLNSPGFPPTQD	540		
Db	481	GVMNRKANRVFGQWVSSSTLWNGEAIKGFVLKCCAYNSLGTSCBTILLNSPGFPPTQD	540		
Qy	541	NISFYATIGVCLLFIIVLTLIICHYKKQPYESQLOMVOVTGSDNRYFYVDREYED	600		
Db	541	NISFYATIGVCLLFIIVLTLIICHYKKQPYESQLOMVOVTGSDNRYFYVDREYED	600		
Qy	601	LKWEFFPRENLEFGKVLGSGAFGKWNATAYGISTGVSIOVAVKMLKEKADSSREALMS	660		
Db	601	LKWEFFPRENLEFGKVLGSGAFGKWNATAYGISTGVSIOVAVKMLKEKADSSREALMS	660		
Qy	661	ELKMTQLGSHENIVNLGACTLSGPIYLIPEYCCYGDLLNVLRSKREKFRHTWTEIPKE	720		
Db	661	ELKMTQLGSHENIVNLGACTLSGPIYLIPEYCCYGDLLNVLRSKREKFRHTWTEIPKE	720		
Qy	721	HNFSFYPTFQSHPNSSMGPSREVQIHPSDQISGLHGNSFSEDEBIYENOKRLEEDL	780		
Db	721	HNFSFYPTFQSHPNSSMGPSREVQIHPSDQISGLHGNSFSEDEBIYENOKRLEEDL	780		
Qy	781	NVLTFEDLLCFAYOVAKGMEFLEPKSCVHRDLAARNVLVTHGKVVKICDFGLARDINSDS	840		
Db	781	NVLTFEDLLCFAYOVAKGMEFLEPKSCVHRDLAARNVLVTHGKVVKICDFGLARDINSDS	840		
Qy	841	NYVVRGNARLPVKWMAPESLFEGYITIKSDVMSYGILLWEIFSLGVNYPGIPVDANFYK	900		
Db	841	NYVVRGNARLPVKWMAPESLFEGYITIKSDVMSYGILLWEIFSLGVNYPGIPVDANFYK	900		
Qy	901	LIQNGFKMDQFPYATEEYIIIMQSCWAFDSRKPSFNLTSFLGCOLADAEAMYQNVDG	960		
Db	901	LIQNGFKMDQFPYATEEYIIIMQSCWAFDSRKPSFNLTSFLGCOLADAEAMYQNVDG	960		
Qy	961	RVSECPHTYQNRPRPSREMDLGLSLPQAQVEDS	993		
Db	961	RVSECPHTYQNRPRPSREMDLGLSLPQAQVEDS	993		

RESULT 7	
RAE25819	
ID	AAE25819 standard; Protein; 993 AA.
XX	
XX	AAE25819;
XX	
DT	15-NOV-2002 (first entry)
XX	
DB	Human receptor protein tyrosine kinase, flk-2.
XX	
KW	Human; receptor protein tyrosine kinase; pTK; haematopoietic cell;
KW	growth factor; therapeutic; macrocytic anaemia; aplastic anaemia;
KW	bone marrow damage; cancer; chemotherapy; radiation; osteopathic;
KW	flk-2 protein.
XX	
OS	Homo sapiens.
XX	
Key	Location/Qualifiers
FT	Peptide
FT	1..27
FT	/label= Signal-peptide
FT	Protein
FT	28..993
FT	/note= "Human mature flk-2 protein"
FT	Domain
FT	1..516
FT	/note= "Extracellular domain"
FT	Domain
FT	517..536
FT	/note= "Transmembrane domain"
FT	Domain
FT	537..966
FT	/note= "Intracellular domain"
XX	
XX	US2002072077-A1.
XX	
XX	13-JUN-2002.
XX	
XX	31-JUL-2001; 2001US-0919408.
XX	
XX	19-NOV-1992; 92US-0977451.
PR	30-APR-1993; 93US-0055289.
PR	31-OCT-1994; 94US-0252498.
PR	15-FEB-1996; 96US-0601891.
PR	10-FEB-1998; 98US-0021324.
PR	10-DEC-1998; 98US-0208786.
PR	02-APR-1991; 91US-0679666.
PR	28-JUN-1991; 91US-0728913.
PR	15-NOV-1991; 91US-0793065.
PR	24-DEC-1991; 91US-0813593.
XX	
XX	(LEMI/) LEMISCHKA I R.
XX	
PI	Lemischka IR;
XX	
XX	WPI; 2002-607237/65.
DR	N-PSDB; AAD42484.
XX	
XX	New protein tyrosine kinase expressed in primitive hematopoietic cells
PT	(HC) and not expressed in mature HC, and ligands for the protein, for
PT	stimulating proliferation of primitive hematopoietic stem cells -
XX	
PS	Claim 37; Page 25-28; 64pp; English.
XX	
CC	The present invention relates to receptor protein tyrosine kinase (pTK)
CC	expressed in primitive haematopoietic cells and not expressed in mature
CC	haematopoietic cells, polynucleotides encoding such proteins and ligands
CC	for the protein. Ligands which bind to pTK of the invention are useful
CC	for stimulating the proliferation and/or differentiation of primitive
CC	mammalian haematopoietic stem cells. The receptor pTK molecules are
CC	useful for stimulating the self renewal of the totipotent haematopoietic
CC	stem cell and to stimulate the development of all cells of haematopoietic
CC	system both in vitro and in vivo. The ligands for the receptors act as
CC	haematopoietic growth factors. The ability of the ligands to stimulate
CC	proliferation of stem cells both in vitro and in vivo has important
CC	therapeutic applications such as treating humans whose primitive stem
CC	cells do not sufficiently undergo self-renewal. It is also useful in
CC	conditions that occur when defects in haematopoietic stem cells or their

CC related growth factors depress the number of white blood cells such as macrocytic and aplastic anaemia or bone marrow damage resulting from cancer chemotherapy and radiation. The present sequence is human receptor pTK, flk-2.		Query Match		100.0%; Score 5274; DB 23; Length 993;	
CC		Best Local Similarity		100.0%; Pred. No. 0;	
CC		Matches 993; Conservative		0; Mismatches 0; Indels 0; Gaps 0;	
Qy	1	MPALARDAGTGLVVFAMIFGTTTQDLPVVKVNLHNKNDSSVGKSSSYPMVSESP	60		
Db	1	MPALARDAGTGLVVFAMIFGTTTQDLPVVKVNLHNKNDSSVGKSSSYPMVSESP	60		
Qy	61	EDLGCALRPQSSGTYEAAAEVDVSASITLQVLVDAPGNISCLWVFKHSSLNCPHFDL	120		
Db	61	EDLGCALRPQSSGTYEAAAEVDVSASITLQVLVDAPGNISCLWVFKHSSLNCPHFDL	120		
Qy	121	QNRGVSVMLKMTQAGEYLLFTQSEATNTYILFTVSRNTLLYTLRRPYFRKMNQD	180		
Db	121	QNRGVSVMLKMTQAGEYLLFTQSEATNTYILFTVSRNTLLYTLRRPYFRKMNQD	180		
Qy	181	ALVCISESVPEPIVWNLCDSGESCKEESPAVVKKEKVLHELFGTDIRCCARNELGRE	240		
Db	181	ALVCISESVPEPIVWNLCDSGESCKEESPAVVKKEKVLHELFGTDIRCCARNELGRE	240		
Qy	241	CTRLEFTIDNLTQPTTLQPLFKVGEPLWIRKAVVNHGFLTWELNKALEEGNYEM	300		
Db	241	CTRLEFTIDNLTQPTTLQPLFKVGEPLWIRKAVVNHGFLTWELNKALEEGNYEM	300		
Qy	301	STYSTNRTMIRLFAFVSSVARNDGYTCSSSKHPSQSALVTIIVKGFINATNSEDYE	360		
Db	301	STYSTNRTMIRLFAFVSSVARNDGYTCSSSKHPSQSALVTIIVKGFINATNSEDYE	360		
Qy	361	IDOHEEFCFVRFKAYPOIRCTWTSRSPFCEQKGLDNGYSISKFCNHKHPGYIFHA	420		
Db	361	IDOHEEFCFVRFKAYPOIRCTWTSRSPFCEQKGLDNGYSISKFCNHKHPGYIFHA	420		
Qy	421	ENDDAQFTKMTFLNRRKQVLAESAASQSCFSDGYPLPSMTWKCKSPNCTEEITE	480		
Db	421	ENDDAQFTKMTFLNRRKQVLAESAASQSCFSDGYPLPSMTWKCKSPNCTEEITE	480		
Qy	481	GVNKRKANRVFGQWSSSTLNMSBAIKGLFKCCAYNSLGTSCETILLNSPGPPFFIQD	540		
Db	481	GVNKRKANRVFGQWSSSTLNMSBAIKGLFKCCAYNSLGTSCETILLNSPGPPFFIQD	540		
Qy	541	NISFYATIGVCLLFIWLTLLI CHYKQKQRYESQLOMVQVGTSSDNEYFYVDREYEYD	600		
Db	541	NISFYATIGVCLLFIWLTLLI CHYKQKQRYESQLOMVQVGTSSDNEYFYVDREYEYD	600		
Qy	601	LKWEFPRENLEFGKVLGSGAFGKVNATAYGISTGVSIOQAVKMLKEKADSSEREA	660		
Db	601	LKWEFPRENLEFGKVLGSGAFGKVNATAYGISTGVSIOQAVKMLKEKADSSEREA	660		
Qy	661	ELKMTTQLGSHENIVNLLGACTLSGPIYLIIFYCCYGDLLNLSRKREKPHRTWTI	720		
Db	661	ELKMTTQLGSHENIVNLLGACTLSGPIYLIIFYCCYGDLLNLSRKREKPHRTWTI	720		
Qy	721	HNFSYPTQSHPNMSPGSRVQIHPDSDQISGLHNSFHSDETEYENQRLSEEDL	780		
Db	721	HNFSYPTQSHPNMSPGSRVQIHPDSDQISGLHNSFHSDETEYENQRLSEEDL	780		
Qy	781	NVLTTEDLLCFAYQVAKGMEFELEKSCVRDLAARNVLTHGVKVKICDFGLARDIMS	840		
Db	781	NVLTTEDLLCFAYQVAKGMEFELEKSCVRDLAARNVLTHGVKVKICDFGLARDIMS	840		
Qy	841	NYVVRGNARLPVKWAPESLFGIYTIKSDVMSYGILLWEIFSLGVNYPGIPVDANFYK	900		
Db	841	NYVVRGNARLPVKWAPESLFGIYTIKSDVMSYGILLWEIFSLGVNYPGIPVDANFYK	900		
Qy	901	LIONGFMDQPPYATEEIIYIMQSCWAFDSRKRSPFNLTSLFGCOLADAEAMYNVDG	960		

Db	901	LIONGFMDQPPYATEEIIYIMQSCWAFDSRKRSPFNLTSLFGCOLADAEAMYNVDG	960		
Qy	961	RVSECPHTYQNRPPSPREMDLGLLSPOAQVEDS	993		
Db	961	RVSECPHTYQNRPPSPREMDLGLLSPOAQVEDS	993		
RESULT 8					
ID	AAR37503	standard; Protein; 993 AA.			
XX	AAR37503;				
XX	25-MAR-2003	(updated)			
DT	19-OCT-1993	(first entry)			
XX	Human flk-2.				
XX	Murine; receptor; protein; tyrosine kinase; pTK; flk-2; primitive;				
KW	hematopoietic cell; mature; family; conserved; region;				
KW	catalytic domain; c-kit; fetal liver kinase; flk; fetal; spleen;				
KW	thymus; adult; brain; bone marrow; multipotential; CFU-Blast colony;				
KW	hierarchy; transduction; T-lymphoid; lineage.				
OS	Homo sapiens.				
XX					
Key	Location/Qualifiers				
FT	1..27				
FT	/note= "Hydrophobic leader sequence"				
FT	28..993				
FT	/note= "Mature murine flk-2"				
FT	28..543				
FT	/note= "Extracellular receptor domain"				
FT	544..563				
FT	/note= "Transmembrane region"				
FT	564..983				
FT	/note= "Intracellular catalytic domain"				
XX					
PN	WO9310136-A1.				
XX					
PD	27-MAY-1993.				
XX					
PF	16-NOV-1992; 92WO-US09893.				
XX					
PR	15-NOV-1991; 91US-0793065.				
XX					
PA	(UYPR-) UNIV PRINCETON.				
XX					
PI	Lemischka IR;				
XX					
DR	WPI; 1993-182479/22.				
DR	N-PSDB; AAQ40915.				
XX					
PT	Totipotent haematopoietic stem cell receptors, their ligands and				
PT	DNA sequences - for treating anaemia(s) and bone marrow damage				
PT	due to e.g. cancer chemotherapy or radiotherapy				
XX					
PS	Claim 39; Fig 1b; 127pp; English.				
XX					
CC	This sequence represents the human receptor protein tyrosine kinase				
CC	(pTK), flk-2. The nucleic acid encoding this receptor is expressed				
CC	in primitive hematopoietic cells and not in mature hematopoietic				
CC	cells. Members of this family of pTK's can be recognised by the				
CC	conserved amino acid regions in the catalytic domain. This family				
CC	of pTK's also contains c-kit. These new receptors are termed fetal				
CC	liver kinases (flk's) after the tissue in which they were discovered.				
CC	flk-2 is also expressed in fetal spleen, fetal thymus, adult brain				
CC	and adult bone marrow. flk-2 is expressed in individual multipotential				
CC	CFU-Blast colonies capable of generating numerous multilineage colonies				
CC	upon replating. It is likely therefore, that flk-2 is expressed in				
CC	the entire primitive portion of the hematopoietic hierarchy. This is				
CC	consistent with flk-2 being important in transducing putative self-				
CC	renewal signals from the environment. flk-2 is the first receptor				

CC pTK known to be expressed in the T-lymphoid lineage.
CC (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 993 AA;

Query Match		99.9%; Score 5271; DB 14; Length 993;
Best Local Similarity		99.9%; Pred. No. 0;
Matches 992; Conservative 1; Mismatches 0; Indels 0; Gaps 0;		
QY	1	MPALARDAGTVPLLVVFSAMIFGTTQDLPVVKVNLHKNNDSSVCKSSYPWVSESP 60
DB	1	MPALARDAGTVPLLVVFSAMIFGTTQDLPVVKVNLHKNNDSSVCKSSYPWVSESP 60
QY	61	EDLGCALRPOSSGTVYAAAVEVDVSASITQLVLDAPGNISCLVFKHSSLNCPHFDL 120
DB	61	EDLGCALRPOSSGTVYAAAVEVDVSASITQLVLDAPGNISCLVFKHSSLNCPHFDL 120
QY	121	QNRGVSWILKMTETQAGEYLLFIQSEATNTYTLFTVSINTLYTLRRPYFRMENQD 180
DB	121	QNRGVSWILKMTETQAGEYLLFIQSEATNTYTLFTVSINTLYTLRRPYFRMENQD 180
QY	181	ALVCISESVPEPIVEWVLCDSQGESCKEESPAVYKKEKVLHFGTDIRCCARNELGRE 240
DB	181	ALVCISESVPEPIVEWVLCDSQGESCKEESPAVYKKEKVLHFGTDIRCCARNELGRE 240
QY	241	CTRLFTIDLNQTPOTTLPLQFLKVGEPILWIRCKAVHNHGFGLTWELNKALEBGNFEM 300
DB	241	CTRLFTIDLNQTPOTTLPLQFLKVGEPILWIRCKAVHNHGFGLTWELNKALEBGNFEM 300
QY	301	STYSTNRMTIRILFAFVSSVARNDTYTTCSSSKHPSQSALVTIVGKGFNATNSSDEYE 360
DB	301	STYSTNRMTIRILFAFVSSVARNDTYTTCSSSKHPSQSALVTIVGKGFNATNSSDEYE 360
QY	361	IDQYEFCEFSVRKAYPOIRCTWTFSPKSPCEQKGLDNGVYSISKPNCHKHQPGEYIFHA 420
DB	361	IDQYEFCEFSVRKAYPOIRCTWTFSPKSPCEQKGLDNGVYSISKPNCHKHQPGEYIFHA 420
QY	421	ENDDAQFTKMTINIRKPKVLAESAASQASCSFGDGYPLSWTKKCDKSPNCTEITE 480
DB	421	ENDDAQFTKMTINIRKPKVLAESAASQASCSFGDGYPLSWTKKCDKSPNCTEITE 480
QY	481	GVNRRKANRKFQGVSSSTLNKSEATKGLFKCCAYNSLGTSCETILLNSPGFPFIQD 540
DB	481	GVNRRKANRKFQGVSSSTLNKSEATKGLFKCCAYNSLGTSCETILLNSPGFPFIQD 540
QY	541	NISFYATIGVCLLPIVVLTLIICHYKKQPRYESQLQWQVGTSSDNEYFYVDFREYED 600
DB	541	NISFYATIGVCLLPIVVLTLIICHYKKQPRYESQLQWQVGTSSDNEYFYVDFREYED 600
QY	601	LKWEFFRENLEFCVKLGSGAFGKMNATAYGISTGVSIOVAVKMLKEKADSSEREAALMS 660
DB	601	LKWEFFRENLEFCVKLGSGAFGKMNATAYGISTGVSIOVAVKMLKEKADSSEREAALMS 660
QY	661	ELKQMTQLGSHENIVNLLGACTLSGPIYLLIFCYCCYDGLNLYRSKREKPHRTWTTEIFE 720
DB	661	ELKQMTQLGSHENIVNLLGACTLSGPIYLLIFCYCCYDGLNLYRSKREKPHRTWTTEIFE 720
QY	721	HNFSFYPTQSHNSMPGSRVQIHPDSDQISGLHGNFSHSEDEIYEYENQKLEBEEDL 780
DB	721	HNFSFYPTQSHNSMPGSRVQIHPDSDQISGLHGNFSHSEDEIYEYENQKLEBEEDL 780
QY	781	NVLTFEDLLCFAYQVAKGMEFLFKCVHRDLAARNVLTGHVKVICDFGLARDIMSDS 840
DB	781	NVLTFEDLLCFAYQVAKGMEFLFKCVHRDLAARNVLTGHVKVICDFGLARDIMSDS 840
QY	841	NYVVRGNARLPVKWMAPELIFEGITYIKSDVWSYGILLMEIFSLGVPYPGIIPVDANFYK 900
DB	841	NYVVRGNARLPVKWMAPELIFEGITYIKSDVWSYGILLMEIFSLGVPYPGIIPVDANFYK 900
QY	901	LQNGFKMDQPPFYATEEIIYIMQSWAFDSRKPSPFNLTSLFLGCQLADAEAEAMYNQVNDG 960
DB	901	LQNGFKMDQPPFYATEEIIYIMQSWAFDSRKPSPFNLTSLFLGCQLADAEAEAMYNQVNDG 960

Query Match

99.8%; Score 5266; DB 14; Length 993;

QY	961	RVSECPHTYQNRPPFSREMDLGLLSPOAQVEDS 993
DB	961	RVSECPHTYQNRPPFSREMDLGLLSPOAQVEDS 993
RESULT 9		
AAR44995		
ID	AAR44995	standard; Protein; 993 AA.
XX	AAR44995;	
XX	25-MAR-2003	(updated)
DT	27-JUN-1994	(first entry)
XX	Human flk-2 receptor protein tyrosine kinase.	
XX	Receptor protein tyrosine kinase; pTK family; foetal liver kinase;	
KW	hflk; primitive; totipotent; haematopoietic cell; stem cell;	
KW	proliferation; stromal cell.	
XX	Homo sapiens.	
OS		
XX	Key	Location/Qualifiers
FT	Peptide	1..27
FT		/label= signal sequence
FT		/notes= "hydrophobic leader"
FT	Protein	28..993
FT		/label= flk-2
FT		/note= "mature protein"
FT	Domain	28..544
FT		/label= extracellular_domain
FT	Region	545..563
FT		/label= transmembrane_region
FT	Domain	564..993
FT		/label= intracellular_catalytic_domain
XX	US5270458-A.	
XX	14-DEC-1993.	
XX	19-NOV-1992;	92US-0977451.
XX	02-APR-1991;	91US-0679666.
PR	28-JUN-1991;	91US-0728913.
PR	15-NOV-1991;	91US-0793065.
PR	24-DEC-1991;	91US-0813593.
PR	26-JUN-1992;	92US-0906397.
PR	12-NOV-1992;	92US-0975049.
PR	19-NOV-1992;	92US-0977451.
XX	(UYPR-) UNIV PRINCETON.	
PA		
XX	Lemischka IR;	
XX	WPI; 1993-405021/50.	
DR	N-PSDB; AAQ53503.	
XX	Isolated nucleic acid molecules of hematopoietic stem cell	
PT	receptor flk-2 - encoding mammalian receptor protein tyrosine	
PT	kinases expressed in primitive haematopoietic cells	
XX	Disclosure; Fig 1b; 60pp; English.	
XX	Nucleic acid sequences coding for murine flk-2 and specified	
CC	subfragments of it are claimed. The human flk-2 coding sequence	
CC	(i.e. AAQ53503) is also disclosed. The murine and human flk-2	
CC	polypeptides represent a new class of receptor protein tyrosine	
CC	kinases which are expressed only in primitive haematopoietic cells.	
CC	(Updated on 25-MAR-2003 to correct PF field.)	
XX	Sequence	993 AA;
SQ		

Best Local Similarity 99.9%; Pred. No. 0;			
Matches 992; Conservative 0; Mismatches 1; Indels 0; Gaps 0;			
QY	1	MPALARDAGTVP	LLVVFSA
Db	1	MPALARDAGTVP	LLVVFSA
QY	61	EDLGCALRPQSSG	TVYEA
Db	61	EDLGCALRPQSSG	TVYEA
QY	121	QNRGVVSMVLK	MTETQAGEY
Db	121	QNRGVVSMVLK	MTETQAGEY
QY	181	ALVCISESVPE	PIVWVLCDSQ
Db	181	ALVCISESVPE	PIVWVLCDSQ
QY	241	CTRLFTIDL	NOTPQTLP
Db	241	CTRLFTIDL	NOTPQTLP
QY	301	STYSTNR	TMRILFAFVSS
Db	301	STYSTNR	TMRILFAFVSS
QY	361	IDQYBEFC	SVRPFKAYPQ
Db	361	IDQYBEFC	SVRPFKAYPQ
QY	421	ENDDAQ	FTKXFTLNIR
Db	421	ENDDAQ	FTKXFTLNIR
QY	481	GYWNRKAN	KVFQNTYSS
Db	481	GYWNRKAN	KVFQNTYSS
QY	541	NTSFYATIG	VCLLFIVVLT
Db	541	NTSFYATIG	VCLLFIVVLT
QY	601	LKWEFPRE	NLEFGKVLG
Db	601	LKWEFPRE	NLEFGKVLG
QY	661	ELKMMTQ	LGSHEINVL
Db	661	ELKMMTQ	LGSHEINVL
QY	721	HNFSFYPT	QSHNSNMP
Db	721	HNFSFYPT	QSHNSNMP
QY	781	NVLTFED	LICFAYQV
Db	781	NVLTFED	LICFAYQV
QY	841	NYVVRGN	ARLPVKWMA
Db	841	NYVVRGN	ARLPVKWMA
QY	901	LIQNGF	KMDQPFYAT
Db	901	LIQNGF	KMDQPFYAT
QY	961	RVSECP	HTYQNR
Db	961	RVSECP	HTYQNR

AAR75961			
ID	AAR75961	standard; Protein; 993 AA.	
XX	AC	AAR75961;	
XX	DT	25-MAR-2003 (updated)	
XX	DT	29-DEC-1995 (first entry)	
XX	DE	Human STK-1.	
XX	KW	STK-1; receptor PTK; protein tyrosine kinase.	
XX	OS	Homo sapiens.	
XX	FT	Key	Location/Qualifiers
XX	FT	Peptide	1..23
XX	FT	/label= signal	
XX	FT	/note= "unique PTK receptor sequence"	
XX	FT	Cleavage-site	25..26
XX	FT	Domain	35..524
XX	FT	/note= "part of extracellular domain contg. 22 Cys residues marked on Fig. 1"	
XX	FT	Modified-site	43..45
XX	FT	/label= potential N-linked glycosylation site	
XX	FT	Modified-site	100..102
XX	FT	/label= see above	
XX	FT	Modified-site	151..153
XX	FT	/label= see above	
XX	FT	Modified-site	250..252
XX	FT	/label= see above	
XX	FT	Modified-site	306..308
XX	FT	/label= see above	
XX	FT	Modified-site	323..325
XX	FT	/label= see above	
XX	FT	Modified-site	351..353
XX	FT	/label= see above	
XX	FT	Modified-site	354..356
XX	FT	/label= see above	
XX	FT	Modified-site	473..475
XX	FT	/label= see above	
XX	FT	Modified-site	502..504
XX	FT	/label= see above	
XX	FT	Region	542..562
XX	FT	/label= transmembrane spanning region	
XX	FT	Region	708..782
XX	FT	/label= kinase insert region	
XX	FT	Domain	617..622
XX	FT	/label= ATP binding domain	
XX	FT	Domain	835..840
XX	FT	/label= WWAPES motifs	
XX	FT	/note= "cytoplasmic domain"	
XX	FT	Peptide	808..813
XX	FT	/note= "used to design PCR oligos"	
XX	FT	Peptide	870..875
XX	FT	/note= "used to design PCR oligos"	
XX	XX	W09519175-A1.	
XX	XX	20-JUL-1995.	
XX	XX	06-JAN-1995;	95WO-US00176.
XX	XX	14-JAN-1994;	94US-0183211.
XX	XX	(UJJO) UNIV JOHNS HOPKINS.	
XX	XX	(UYPE-) UNIV PENNSYLVANIA.	
XX	XX	Civin CI, Gewirtz AM, Small D;	
XX	XX	WPI; 1995-263709/34.	
XX	XX	N-PSDB; AAQ91356.	

PT used to treat neoplastic diseases and as bone marrow purging agents for
 PT treating leukaemia and neoplasia
 PS Disclosure; Fig 1; 66pp; English.
 XX

CC The STK-1 gene encodes a receptor PTK which is expressed in
 CC proliferating hematopoietic stem cells but not in quiescent stem
 CC cells. The STK-1 gene is also expressed in certain malignant cells
 CC of non-hematopoietic origin. An antisense oligo specific for STK-1
 CC is an oligo having a sequence (i) capable of forming a stable
 CC triplex with a portion of the STK-1 gene, or (ii) capable of forming
 CC a stable duplex with a portion of an mRNA transcript of the STK-1
 CC gene. Antisense oligos capable of forming a stable duplex with a
 CC portion of a STK-1 mRNA transcript are given in AAQ91536 FT and in
 CC AAQ91537 and AAQ91538. The antisense oligos of the invention are useful
 CC in the treatment of hematologic malignancies characterized by
 CC STK-1 expression. Several of the conserved domains of PTKs
 CC including the ATP binding domain and the WNAPEs motifs are
 CC found in the STK-1 protein (see AAR75961 FT).
 CC (Updated on 25-MAR-2003 to correct PN field.)
 XX Sequence 993 AA;

Query Match 99.8%; Score 5266; DB 16; Length 993;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 992; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MPALARDAGTVPVLLVVSAMIFGTITNODLPVTKVLIHKNNDSSVGKSSSPMVSESP 60
 DB 1 MPALARDAGTVPVLLVVSAMIFGTITNODLPVTKVLIHKNNDSSVGKSSSPMVSESP 60

QY 61 EDLGCALRPQSSGTVEAAAVEVDVSGASITLOVLVDAPGNISCLVWFKHSNLCQPHFDL 120
 DB 61 EDLGCALRPQSSGTVEAAAVEVDVSGASITLOVLVDAPGNISCLVWFKHSNLCQPHFDL 120

QY 121 QNRGVVSMVLKMTQAGEYLLFIOSEATNYTILFTVSRNTLLYLRLRPYFRKMNQD 180
 DB 121 QNRGVVSMVLKMTQAGEYLLFIOSEATNYTILFTVSRNTLLYLRLRPYFRKMNQD 180

QY 181 ALVCISESVEPIVEWVLCDSQGESCKERSPAVVKKEKVLHFLGTDTRCCARNEIGRE 240
 DB 181 ALVCISESVEPIVEWVLCDSQGESCKERSPAVVKKEKVLHFLGTDTRCCARNEIGRE 240

QY 241 CTRLFTIDLNQPTTLPOLFLKVGELPLTRCKAVHNVHGFGLTWELNKALEEGNYFEM 300
 DB 241 CTRLFTIDLNQPTTLPOLFLKVGELPLTRCKAVHNVHGFGLTWELNKALEEGNYFEM 300

QY 301 STYSTNRMTIRILFAFVSSVARNDTGYTCSSSKHPSQSALVTIVGKFINATNSSEYD 360
 DB 301 STYSTNRMTIRILFAFVSSVARNDTGYTCSSSKHPSQSALVTIVGKFINATNSSEYD 360

QY 361 IDQYEEFCFSVRKAYPQIRCTWTWFSRKPFPCEQKGLDNGYSISKFCNKHQPGEYIFHA 420
 DB 361 IDQYEEFCFSVRKAYPQIRCTWTWFSRKPFPCEQKGLDNGYSISKFCNKHQPGEYIFHA 420

QY 421 ENDDAQFTKMTLNIRRKQVLAESAASQSCFSDGYPLPSWTWKKCSKSPNCTBEITE 480
 DB 421 ENDDAQFTKMTLNIRRKQVLAESAASQSCFSDGYPLPSWTWKKCSKSPNCTBEITE 480

QY 481 GWNKRNKRVKQVSSSTLNNSEAIKGLVKCCAYNSLGTSCETILLNSGPPFPFIOD 540
 DB 481 GWNKRNKRVKQVSSSTLNNSEAIKGLVKCCAYNSLGTSCETILLNSGPPFPFIOD 540

QY 541 NISFYATIGVCLLFIIVLTLCHYKKQFRYESQLQVQVVTGSSDNEYFYVDREYED 600
 DB 541 NISFYATIGVCLLFIIVLTLCHYKKQFRYESQLQVQVVTGSSDNEYFYVDREYED 600

QY 601 LKWEPPRENLEKVLGSGAFKGMNATAYGSKTGVSIOVAVKMLKEKADSSEREAAMS 660
 DB 601 LKWEPPRENLEKVLGSGAFKGMNATAYGSKTGVSIOVAVKMLKEKADSSEREAAMS 660

QY 661 ELKMMTOLGSHENIVNLGACTLSGPIYLIFCYCCYDGLLNLRSKREKFRHTWTWEIFKE 720

DB 661 ELKMMTOLGSHENIVNLGACTLSGPIYLIFCYCCYDGLLNLRSKREKFRHTWTWEIFKE 720
 QY 721 HNFSPYPTFQSHNSNMPGSRREVQIHPDSQISGLHNSFHSDEIEYENOKRLEEDL 780
 DB 721 HNFSPYPTFQSHNSNMPGSRREVQIHPDSQISGLHNSFHSDEIEYENOKRLEEDL 780
 QY 781 NVLTFEDLLCFAYQVAKGMFLFKSCVHRDLAARNVLVTHGKVKIKCDFGLARDIMSDS 840
 DB 781 NVLTFEDLLCFAYQVAKGMFLFKSCVHRDLAARNVLVTHGKVKIKCDFGLARDIMSDS 840
 QY 841 NYVVRGNARLPVKWMAPESELFEGITTIKSDVMSYGILLWEIFSLGVNPPYGPVPDANFYK 900
 DB 841 NYVVRGNARLPVKWMAPESELFEGITTIKSDVMSYGILLWEIFSLGVNPPYGPVPDANFYK 900
 QY 901 LIQNGFKMDOPFYATBEIYIIMQSCWAFDSRKRPSFNNLTSFLGCQLADAEEMVQNVGD 960
 DB 901 LIQNGFKMDOPFYATBEIYIIMQSCWAFDSRKRPSFNNLTSFLGCQLADAEEMVQNVGD 960
 QY 961 RVSECPHTYQNRPPFSREMDLGLLSPQAQVEDS 993
 DB 961 RVSECPHTYQNRPPFSREMDLGLLSPQAQVEDS 993

RESULT 11

AAR81869
 ID AAR81869 standard; Protein; 993 AA.
 XX
 AC AAR81869;
 XX
 DT 19-MAR-1996 (first entry)
 XX
 DE Human Flk2/flt3 tyrosine kinase receptor.
 XX
 KW Tyrosine kinase receptor; flk2; flt3; agonist; monoclonal antibody;
 KW haematopoiesis; hypoplasia; anaemia; thrombocytopenia; stem cell.
 XX
 OS Homo sapiens.
 XX
 FN W09527062-A1.
 XX
 PD 12-OCT-1995.
 XX
 PF 23-MAR-1995; 95WO-US03718.
 XX
 PR 04-APR-1994; 94US-0222299.
 XX
 PA (GETH) GENENTECH INC.
 XX
 PI Bennett BD, Broz SD, Matthews W, Zeigler PC;
 XX
 DR WPI; 1995-358636/46.
 DR N-P8DB; AAT00802.
 XX
 PT Agonist antibody which binds to flk2-flt3 tyrosine kinase receptor
 PT - enhances proliferation of haematopoietic stem cells, in the
 PT treatment of hypoplasia, anaemia, etc.
 XX
 PS Disclosure; Page 44-47; 59pp; English.
 XX
 CC DNA (AAT00802) coding for murine stem cell tyrosine kinase receptor flk2
 CC (also called flt3) (AAR81869) was obtd. by RT-PCR amplification of RNA
 CC isolated from mid-gestation mouse foetal livers using primers based
 CC on the murine flt3 sequence, and subcloning of the product into pRK5.1.
 CC An flk2/flt3 extracellular domain-IgG1 Fc fusion protein was
 CC constructed and used to raise agonist antibodies able to bind to, and
 CC activate, flk2/flt3. The human flk2/flt3 receptor gene (AAT00802) and
 CC amino acid sequence (AAR81869) are also given.
 XX
 SQ Sequence 993 AA;

Query Match 99.8%; Score 5265; DB 16; Length 993;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 992; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MPALARDAGTVPLLVVFSAMIFGTITNODLPVKVLINHKNDSSVGKSSYPWVSESP 60
DB 1 MPALARDAGTVPLLVVFSAMIFGTITNODLPVKVLINHKNDSSVGKSSYPWVSESP 60
QY 61 EDLGCALRPOSSGTYEAAAEVVDVSASITLQVLDAPGNISCLWVKHSSLNCOHPDL 120
DB 61 EDLGCALRPOSSGTYEAAAEVVDVSASITLQVLDAPGNISCLWVKHSSLNCOHPDL 120
QY 121 QNRGVSVILKMTETQAGEYLLFQSEATNTILFTVSIRNTLYTLRPPFRKXENQD 180
DB 121 QNRGVSVILKMTETQAGEYLLFQSEATNTILFTVSIRNTLYTLRPPFRKXENQD 180
QY 181 ALVCISESVPEPIVEMVLCDSSQESCKEESPAVVKKEKVLHELFGTDIRCCARNELGRE 240
DB 181 ALVCISESVPEPIVEMVLCDSSQESCKEESPAVVKKEKVLHELFGTDIRCCARNELGRE 240
QY 241 CTRLFTIDLNQTPOTTLPOLFLKVGEPILWIRKAVHVNHGFLTWELNKALEEGNYFEM 300
DB 241 CTRLFTIDLNQTPOTTLPOLFLKVGEPILWIRKAVHVNHGFLTWELNKALEEGNYFEM 300
QY 301 STYSTNRMTIRILFAFVSVARNDTGYYTCSSSKHPSQSALVTIIVGKGFINATNSSEYD 360
DB 301 STYSTNRMTIRILFAFVSVARNDTGYYTCSSSKHPSQSALVTIIVGKGFINATNSSEYD 360
QY 361 IDQYEEFCFSVRFKAYPOIRCTWTFSRKSPCEQKGLDNGYSISKFCNKHQPGEYIFHA 420
DB 361 IDQYEEFCFSVRFKAYPOIRCTWTFSRKSPCEQKGLDNGYSISKFCNKHQPGEYIFHA 420
QY 421 ENDDAQFTQMTNIRRKQVLAESAASQASCFSDGYPLPSWTWKKCSKSPNCTEITE 480
DB 421 ENDDAQFTQMTNIRRKQVLAESAASQASCFSDGYPLPSWTWKKCSKSPNCTEITE 480
QY 481 GYWNRRKANRKFQGVSSSTLNANSEAIKGLVKKCAVNSLGTSCETILLNSPGPPFIOD 540
DB 481 GYWNRRKANRKFQGVSSSTLNANSEAIKGLVKKCAVNSLGTSCETILLNSPGPPFIOD 540
QY 541 NISFYATIGVCLLFIIVLTLILCHYKKQFRYESQOLQVQVGTSSDNEYFYVDFREYED 600
DB 541 NISFYATIGVCLLFIIVLTLILCHYKKQFRYESQOLQVQVGTSSDNEYFYVDFREYED 600
QY 601 LKWEPPRENLEFGKVLGSAFGKWMNATAYGISTGVSIOQAVKMLKEKADSSEREAALS 660
DB 601 LKWEPPRENLEFGKVLGSAFGKWMNATAYGISTGVSIOQAVKMLKEKADSSEREAALS 660
QY 661 ELKMTQLGSHENIVNLGACTLSGPYILIFEYCCYGDLLNLYRSKREKPHRTWTIFKE 720
DB 661 ELKMTQLGSHENIVNLGACTLSGPYILIFEYCCYGDLLNLYRSKREKPHRTWTIFKE 720
QY 721 HNFSFYPTQSHNSMPSGSRVQIHPDSDQISGLHGNSFHSDEIEYENQKLEEEEDL 780
DB 721 HNFSFYPTQSHNSMPSGSRVQIHPDSDQISGLHGNSFHSDEIEYENQKLEEEEDL 780
QY 781 NVLTFEDLLCFAYQVAKGMEFLBFKSCVHRDLAARNVLVTHGKVKICDPLGLARDIMSDS 840
DB 781 NVLTFEDLLCFAYQVAKGMEFLBFKSCVHRDLAARNVLVTHGKVKICDPLGLARDIMSDS 840
QY 841 NYVVRGNARLPVKWMAPELFEIGYITKSDVWSVGLLWEIFSIVGNYPYGPVPDANFYK 900
DB 841 NYVVRGNARLPVKWMAPELFEIGYITKSDVWSVGLLWEIFSIVGNYPYGPVPDANFYK 900
QY 901 LIQNGFKMDOPFYATIEIYIMOSCAFDSRKPSPFNLTSLFGCQLADAEAEAMYQNVDG 960
DB 901 LIQNGFKMDOPFYATIEIYIMOSCAFDSRKPSPFNLTSLFGCQLADAEAEAMYQNVDG 960
QY 961 RVSECPHTYQNRPPFSREMDLGLLSPQAQVEDS 993
DB 961 RVSECPHTYQNRPPFSREMDLGLLSPQAQVEDS 993

RESULT 12
ID AAR31376
AAR31376 standard; Protein; 1167 AA.

XX AAR31376;
AC 25-MAR-2003 (updated)
DT 25-JUN-1993 (first entry)
XX Human flk-2.
KW Human; receptor; protein; tyrosine kinase; pTK; primitive; mammalian;
KW hematopoietic cell; pHC; mature; mHC; fetal; liver kinase 2; flk-2;
KW liver; spleen; thymus; adult; brain; marrow; thymocyte; subset;
KW multipotential; T-lymphoid; lineage.
XX Homo sapiens.
OS
FH Key Location/Qualifiers
FT Misc-difference 1013 /note= "Nonsense mutation"
FT Misc-difference 1038 /note= "Nonsense mutation"
FT Misc-difference 1042 /note= "Nonsense mutation"
FT Misc-difference 1057 /note= "Nonsense mutation"
FT Misc-difference 1102 /note= "Nonsense mutation"
FT Misc-difference 1117 /note= "Nonsense mutation"
FT Misc-difference 1127 /note= "Nonsense mutation"
XX WO9300349-A1.
XX 07-JAN-1993.
XX 26-JUN-1992; 92WO-US05401.
XX 28-JUN-1991; 91US-0728913.
PR 15-NOV-1991; 91US-0793065.
PR 24-DEC-1991; 91US-0813593.
PR 02-APR-1992; 92WO-US02750.
XX (UYPR-) UNIV PRINCETON.
XX Lemischka IR;
XX WPI; 1993-036323/04.
XX N-PSDB; AAQ35250.
XX Nucleic acid encoding receptor protein tyrosine kinase - allows
PT development of ligands to stimulate proliferation and/or
PT differentiation of mammalian haematopoietic stem cells
XX
PS Claim 9; Fig 1b; 78pp; English.
XX This sequence represents a human receptor protein tyrosine kinase
CC which belongs to a new functional class of protein tyrosine kinases
CC (pTKs). pTKs in this class are expressed in primitive mammalian
CC hematopoietic (pHC) cells but not in mature hematopoietic cells (mHC).
CC This protein is an example of a receptor pTK and is called fetal liver
CC kinase 2 (flk-2). flk-2 is expressed in fetal liver, spleen and thymus,
CC and adult brain and marrow. Expression of flk-2 mRNA occurs in the
CC most primitive thymocyte subset, which is believed to be uncommitted.
CC Therefore, thymocytes expressing flk-2 may be multipotential. flk-2
CC is the first receptor tyrosine kinase known to be expressed in the
CC T-lymphoid lineage.
XX (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 1167 AA;

Query Match 99.8%; Score 5262; DB 14; Length 1167;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 991; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MPALARDAGTVPVLLVFSAMIFGTTNQDLPVVKVCLNHKNNDSSVGKSSSYPMVSESP 60
 Db 20 MPALARDAGTVPVLLVFSAMIFGTTNQDLPVVKVCLNHKNNDSSVGKSSSYPMVSESP 79
 QY 61 EDLGCALRPOSSGTVYEAADVEVDVSAITLOVLVDAPGNISCLWVFKHSLNCQPHFDL 120
 Db 80 EDLGCALRPOSSGTVYEAADVEVDVSAITLOVLVDAPGNISCLWVFKHSLNCQPHFDL 139
 QY 121 QNRGVSVMLKMTETQAGEYLLFIQSEATNTILFTVSRINTLLYTLRRPYFRKMENOD 180
 Db 140 QNRGVSVMLKMTETQAGEYLLFIQSEATNTILFTVSRINTLLYTLRRPYFRKMENOD 199
 QY 181 ALVCISSEVPPIVEVWVLCDSQGESCKEESPAVVKKEKVLHELFGTDIRCCARNELGRE 240
 Db 200 ALVCISSEVPPIVEVWVLCDSQGESCKEESPAVVKKEKVLHELFGTDIRCCARNELGRE 259
 QY 241 CTRLFTIDLNOTPQTTLPOLFLKVGPELWIRCAVHVNHGFGLTWELNKALBEGNYFEM 300
 Db 260 CTRLFTIDLNOTPQTTLPOLFLKVGPELWIRCAVHVNHGFGLTWELNKALBEGNYFEM 319
 QY 301 STYSTNRTMIRILFAFVSSVARNDTGYTCCSSSKHPSQSALVTIVGKGFINATNSSEDE 360
 Db 320 STYSTNRTMIRILFAFVSSVARNDTGYTCCSSSKHPSQSALVTIVGKGFINATNSSEDE 379
 QY 361 IDQEEFCFVRKAYPOIRCTWTFRKSPPCEQKGLDNGYSISKFNKHQPGVEYIFHA 420
 Db 380 IDQEEFCFVRKAYPOIRCTWTFRKSPPCEQKGLDNGYSISKFNKHQPGVEYIFHA 439
 QY 421 ENDDAQTFMTNIRRKQVLAESAQSCFSDGYPLPSWTWKCKSDKSPNCTBEITE 480
 Db 440 ENDDAQTFMTNIRRKQVLAESAQSCFSDGYPLPSWTWKCKSDKSPNCTBEITE 499
 QY 481 GWNKRNKRVKGVSSSTLNMSEAIGFLKCCAYNSLGTSCETILLNSGPPFPFIOD 540
 Db 500 GWNKRNKRVKGVSSSTLNMSEAIGFLKCCAYNSLGTSCETILLNSGPPFPFIOD 559
 QY 541 NISFYATIGVCLLFTVLLCHYKQFRYESOLOQVQVGTSSDNEYFYVDFREYED 600
 Db 560 NISFYATIGVCLLFTVLLCHYKQFRYESOLOQVQVGTSSDNEYFYVDFREYED 619
 QY 601 LKWEPPRENLEPGKVGSAFGKVMNATAYIGSKTGVSIQVAVKMLKEKADSSERREALMS 660
 Db 620 LKWEPPRENLEPGKVGSAFGKVMNATAYIGSKTGVSIQVAVKMLKEKADSSERREALMS 679
 QY 661 ELKMTOLGSHENIVNLGACTLSGPIYLIIFYCCVGLLNLVLRKREKPHRTWTEIPKE 720
 Db 680 ELKMTOLGSHENIVNLGACTLSGPIYLIIFYCCVGLLNLVLRKREKPHRTWTEIPKE 739
 QY 721 HNFSPYPTFQSHPNSSMFGSREVQIHPDSQDLSGLHNSFHSDEIEYENOKRLEEDL 780
 Db 740 HNFSPYPTFQSHPNSSMFGSREVQIHPDSQDLSGLHNSFHSDEIEYENOKRLEEDL 799
 QY 781 NVLTFEDLLCFAYQVAKMEFLFKSVHRDLAARNVLVTHGKVVKICDFGLARDIMSDS 840
 Db 800 NVLTFEDLLCFAYQVAKMEFLFKSVHRDLAARNVLVTHGKVVKICDFGLARDIMSDS 859
 QY 841 NYVVRGARNLPVKWAPESLPEGIYITKSDVMSYGILLWEIFSLGVNYPGIPVDANFYK 900
 Db 860 NYVVRGARNLPVKWAPESLPEGIYITKSDVMSYGILLWEIFSLGVNYPGIPVDANFYK 919
 QY 901 LIQNGFKMDQFYATEEYIYIMQSCWAFDSKRPSFPNLTSLGCOLADAEAMYQNVGD 960
 Db 920 LIQNGFKMDQFYATEEYIYIMQSCWAFDSKRPSFPNLTSLGCOLADAEAMYQNVGD 979
 QY 961 RVSECPHYQNRPRFREMMDLGLLSPOAQVEDS 993
 Db 980 RVSECPHYQNRPRFREMMDLGLLSPOAQVEDS 1012

RESULT 13
 AAW63588
 ID AAW63588 standard; Protein; 983 AA.

XX AAW63588;
 XX AC
 XX DT 12-OCT-1998 (first entry)
 XX DE Human receptor type protein kinase FLT3 protein SEQ ID NO:19.
 XX DE Human receptor type protein kinase; FLT3; diagnosis; leukaemia;
 XX KW membrane-vicinal region; tyrosine kinase; juxtamembrane region.
 XX OS Homo sapiens.
 XX PN WO9817808-A1.
 XX PD 30-APR-1998.
 XX PF 13-OCT-1997; 97MO-JP03667.
 XX PR 18-OCT-1996; 96JP-0297329.
 XX PA (TAKI) TAKARA SHUZO CO LTD.
 XX PI Yokota S;
 XX DR WPI; 1998-362333/31.
 XX DR N-PSDB; AAV39041.
 XX PT Nucleic acid sequences encoding receptor type protein kinase -
 XX PT useful for diagnosing adult acute leukaemia of, e.g. FAB sub-classes
 XX PS M2
 XX PS Claim 6; Page 45-50; 80pp; Japanese.
 XX CC New nucleic acid sequences have been isolated which encode receptor type
 XX CC protein kinases (especially a tyrosine kinases) having tandem repeats in
 XX CC the juxtamembrane region. Also described in the present invention are:
 XX CC (1) any one of 10 amino acid (aa) sequences (AAW63580 to AAW63589)
 XX CC representing FLT3 juxtamembrane receptor type protein kinases found in
 XX CC leukaemia patients of FAB subclass M2, M4 or M5 (nucleic acid sequences
 XX CC encoding the proteins are given in AAV39028 to AAV39042); (2) antibodies
 XX CC recognising the kinases or their portions including the tandem repeat
 XX CC region; (3) nucleic acid sequences hybridising with the nucleic acids;
 XX CC (4) a method for detecting the nucleic acid sequences in human tissue
 XX CC samples by: (i) extracting nucleic acid; (ii) amplifying the kinase
 XX CC gene, and (iii) determining the size of the gene for comparing the size
 XX CC of the normal gene not containing tandem repeats, and (5) kits for
 XX CC carrying out the detection. The products and methods may be used for
 XX CC diagnosing adult acute leukaemia of FAB subclasses M2, M4 and M5.
 XX SQ Sequence 983 AA;
 Query Match 97.8%; Score 5159; DB 19; Length 983;
 Best Local Similarity 99.0%; Pred. No. 0;
 Matches 973; Conservative 0; Mismatches 0; Indels 10; Gaps 1;
 QY 1 MPALARDAGTVPVLLVFSAMIFGTTNQDLPVVKVCLNHKNNDSSVGKSSSYPMVSESP 60
 Db 1 MPALARDAGTVPVLLVFSAMIFGTTNQDLPVVKVCLNHKNNDSSVGKSSSYPMVSESP 60
 QY 61 EDLGCALRPOSSGTVYEAADVEVDVSAITLOVLVDAPGNISCLWVFKHSLNCQPHFDL 120
 Db 61 EDLGCALRPOSSGTVYEAADVEVDVSAITLOVLVDAPGNISCLWVFKHSLNCQPHFDL 120
 QY 121 QNRGVSVMLKMTETQAGEYLLFIQSEATNTILFTVSRINTLLYTLRRPYFRKMENOD 180
 Db 121 QNRGVSVMLKMTETQAGEYLLFIQSEATNTILFTVSRINTLLYTLRRPYFRKMENOD 180
 QY 181 ALVCISSEVPPIVEVWVLCDSQGESCKEESPAVVKKEKVLHELFGTDIRCCARNELGRE 240
 Db 181 ALVCISSEVPPIVEVWVLCDSQGESCKEESPAVVKKEKVLHELFGTDIRCCARNELGRE 240
 QY 241 CTRLFTIDLNOTPQTTLPOLFLKVGPELWIRCAVHVNHGFGLTWELNKALBEGNYFEM 300
 Db 241 CTRLFTIDLNOTPQTTLPOLFLKVGPELWIRCAVHVNHGFGLTWELNKALBEGNYFEM 300

Db 241 CTRLFTIDLNOTPQTTLPOLFLKVGEPWIRCKAVHVNHGRLTWELNKALEEGNYFEM 300
Qy 301 STYSTNRMTIRILFAFVSSVARNDTGYTCSSSKHPSQSALVTIVGKGFINATNSSEDEYE 360
Db 301 STYSTNRMTIRILFAFVSSVARNDTGYTCSSSKHPSQSALVTIVGKGFINATNSSEDEYE 360
Qy 361 IDQYEEFCFVRKAYPQIRCTWTFSRKSFPCEQKGLDNGYSISKFCNHKHQPGYIFHA 420
Db 361 IDQYEEFCFVRKAYPQIRCTWTFSRKSFPCEQKGLDNGYSISKFCNHKHQPGYIFHA 420
Qy 421 ENDDAQFTKMTLNIRRRKQVLAESAASQSCFSDGYPLPSMTWKCKSDKSPNCTEITE 480
Db 421 ENDDAQFTKMTLNIRRRKQVLAESAASQSCFSDGYPLPSMTWKCKSDKSPNCTEITE 480
Qy 481 GVMNRKANRKFQGVQWSSSTLNMSAIGFLVKCCAYNSLGTSCETILLNSGPPFFIQD 540
Db 481 GVMNRKANRKFQGVQWSSSTLNMSAIGFLVKCCAYNSLGTSCETILLNSGPPFFIQD 540
Qy 541 NISFYATIGVCLLFTIVLTLCHIKYKQPRYESQIQMVQVTGSSDNEYFYVDREYED 600
Db 541 NISFYATIGVCLLFTIVLTLCHIKYKQPRYESQIQMVQVTGSSDNEYFYVDREYED 600
Qy 601 L-----KWEPPRENLEFGKVLGSGAFKVMNATAYGISKTGVSTQVAVKMLKEKA 650
Db 601 LKWEPPRENLEFGKVLGSGAFKVMNATAYGISKTGVSTQVAVKMLKEKA 650
Qy 651 DSSREALMSLKMWTQGLSHENIVNLLGACTLSGPIVILFYCCYGDILLNLSRKREKF 710
Db 651 DSSREALMSLKMWTQGLSHENIVNLLGACTLSGPIVILFYCCYGDILLNLSRKREKF 710
Qy 711 HRTWTEIFKEHNFSPYPTQSHPNSSMPGSRVQIHPSDQISGLHGNFSHSEDEIEYEN 770
Db 721 HRTWTEIFKEHNFSPYPTQSHPNSSMPGSRVQIHPSDQISGLHGNFSHSEDEIEYEN 780
Qy 771 OKRLEEDLNVLTEDELLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVKICDF 830
Db 781 OKRLEEDLNVLTEDELLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVKICDF 840
Qy 831 GLARDIMSDSNVVRGNARLPVKWAPESLFGIYTIKSDVMSYIGILLWEIFSLGVNYP 890
Db 841 GLARDIMSDSNVVRGNARLPVKWAPESLFGIYTIKSDVMSYIGILLWEIFSLGVNYP 900
Qy 891 GIPVDANFYKLIQNGFKMDQPPYATEEIIYIMQSCWAFDSRKRPPNLTSLGCOLADA 950
Db 901 GIPVDANFYKLIQNGFKMDQPPYATEEIIYIMQSCWAFDSRKRPPNLTSLGCOLADA 960
Qy 951 BEAMQYNDGRVSECPHTYQNER 973
Db 961 BEAMQYNDGRVSECPHTYQNER 983

RESULT 14

AAW63587
ID AAW63587 standard; Protein; 986 AA.

XX AC AAW63587;

XX DT 12-OCT-1998 (first entry)

XX DE Human receptor type protein kinase FLT3 protein SEQ ID NO:18.

XX KW Human; receptor type protein kinase; FLT3; diagnosis; leukaemia;
XX KN membrane-vicinal region; tyrosine kinase; juxtamembrane region.

XX OS Homo sapiens.

XX PN WO9817808-A1.

XX PD 30-APR-1998.

XX PF 13-OCT-1997; 97WO-JP03667.

XX PR 18-OCT-1996; 96JP-0297329.

XX (TAKI) TAKARA SHUZO CO LTD.
XX Yokota S;
XX WPI; 1998-362333/31.
XX N-PSDB; AAV39040.
XX Nucleic acid sequences encoding receptor type protein kinase -
XX useful for diagnosing adult acute leukaemia of, e.g. FAB sub-classes
XX M2
XX Claim 6; Page 40-45; 80pp; Japanese.
XX New nucleic acid sequences have been isolated which encode receptor type
XX protein kinases (especially a tyrosine kinases) having tandem repeats in
XX the juxtamembrane region. Also described in the present invention are:
XX (1) any one of 10 amino acid (aa) sequences (AAW63580 to AAW63589)
XX representing FLT3 juxtamembrane receptor type protein kinases found in
XX leukaemia patients of FAB subclass M2, M4 or M5 (nucleic acid sequences
XX encoding the proteins are given in AAV39028 to AAV39042); (2) antibodies
XX recognising the kinases or their portions including the tandem repeat
XX region; (3) nucleic acid sequences hybridising with the nucleic acids;
XX (4) a method for detecting the nucleic acid sequences in human tissue
XX samples by: (i) extracting nucleic acid; (ii) amplifying the kinase
XX gene, and (iii) determining the size of the gene for comparing the size
XX of the normal gene not containing tandem repeats, and (5) kits for
XX carrying out the detection. The products and methods may be used for
XX diagnosing adult acute leukaemia of FAB subclasses M2, M4 and M5.
XX Sequence 986 AA;
SQ Query Match 97.8%; Score 5157.5; DB 19; Length 986;
Best Local Similarity 98.7%; Pred. No. 0;
Matches 973; Conservative 0; Mismatches 0; Indels 13; Gaps 1;
Qy 1 MPALARDAGTVPLLVVFSAAMFGTTITNODLPVICKVLINHKNDSSVGKSSYPMVSESP 60
Db 1 MPALARDAGTVPLLVVFSAAMFGTTITNODLPVICKVLINHKNDSSVGKSSYPMVSESP 60
Qy 61 EDLGCALPQSSGTGYEAAAEVDVSASITLQVLDPAGNISCLWVFKHSSLNCPHFDL 120
Db 61 EDLGCALPQSSGTGYEAAAEVDVSASITLQVLDPAGNISCLWVFKHSSLNCPHFDL 120
Qy 121 QNRGVSVMLKMTQTQAGEYLLFIOSEATNTYTLFTVSRNTLTYTLRRPYFRKMENQD 180
Db 121 QNRGVSVMLKMTQTQAGEYLLFIOSEATNTYTLFTVSRNTLTYTLRRPYFRKMENQD 180
Qy 181 ALVCISESVPEPIVEWVLCDQSQESCKEESPAVKKKEKVLHELFGTDIRCCARNEGRE 240
Db 181 ALVCISESVPEPIVEWVLCDQSQESCKEESPAVKKKEKVLHELFGTDIRCCARNEGRE 240
Qy 241 CTRLFTIDLNOTPQTTLPOLFLKVGEPWIRCKAVHVNHGRLTWELNKALEEGNYFEM 300
Db 241 CTRLFTIDLNOTPQTTLPOLFLKVGEPWIRCKAVHVNHGRLTWELNKALEEGNYFEM 300
Qy 301 STYSTNRMTIRILFAFVSSVARNDTGYTCSSSKHPSQSALVTIVGKGFINATNSSEDEYE 360
Db 301 STYSTNRMTIRILFAFVSSVARNDTGYTCSSSKHPSQSALVTIVGKGFINATNSSEDEYE 360
Qy 361 IDQYEEFCFVRKAYPQIRCTWTFSRKSFPCEQKGLDNGYSISKFCNHKHQPGYIFHA 420
Db 361 IDQYEEFCFVRKAYPQIRCTWTFSRKSFPCEQKGLDNGYSISKFCNHKHQPGYIFHA 420
Qy 421 ENDDAQFTKMTLNIRRRKQVLAESAASQSCFSDGYPLPSMTWKCKSDKSPNCTEITE 480
Db 421 ENDDAQFTKMTLNIRRRKQVLAESAASQSCFSDGYPLPSMTWKCKSDKSPNCTEITE 480
Qy 481 GVMNRKANRKFQGVQWSSSTLNMSAIGFLVKCCAYNSLGTSCETILLNSGPPFFIQD 540
Db 481 GVMNRKANRKFQGVQWSSSTLNMSAIGFLVKCCAYNSLGTSCETILLNSGPPFFIQD 540
Qy 541 NISFYATIGVCLLFTIVLTLCHIKYKQPRYESQIQMVQVTGSSDNEYFYVDREYED 600

Db 541 NISFYATIGVCLLFIVVLTLLCHYKQFRYESQLQWVQVGTSSDNEYFYVDREYEYD 600
QY 593 -----DFREYEDLKWEFFRENLEFGVLSGAFGKVMNATAYGISTGYSIQVAVKMLK 647
Db 601 LKWEFFDREYEDLKWEFFRENLEFGVLSGAFGKVMNATAYGISTGYSIQVAVKMLK 660
QY 648 EKADSSEREALMSLKMWTQLGSHENIVNLLGACTLSGPIYLIPEYCCYGDLLNLYRSKR 707
Db 661 EKADSSEREALMSLKMWTQLGSHENIVNLLGACTLSGPIYLIPEYCCYGDLLNLYRSKR 720
QY 708 EKPHRTWTEIFKEHNFSFYPTFQSHPNSSMPGSRVQIHPDSDQISGLHGNFSHSEDEIE 767
Db 721 EKPHRTWTEIFKEHNFSFYPTFQSHPNSSMPGSRVQIHPDSDQISGLHGNFSHSEDEIE 780
QY 768 YENOKLEEBEDLNVLTFEDLLCFAYQVAKGMFEFLFKSCVHRDLAARNVLVTHGKVVKI 827
Db 781 YENOKLEEBEDLNVLTFEDLLCFAYQVAKGMFEFLFKSCVHRDLAARNVLVTHGKVVKI 840
QY 828 CDFGLARDIMSDSNVVRGNARLPVKWMAPESLFEGIYTIKSDVWSYGILLWEIFSLGVN 887
Db 841 CDFGLARDIMSDSNVVRGNARLPVKWMAPESLFEGIYTIKSDVWSYGILLWEIFSLGVN 900
QY 888 PYPGIPVDANFYKLIQNGFKMDQFPYATEBIYIIMQSCWAFDSKRKPSFPNLTSFLGCQL 947
Db 901 PYPGIPVDANFYKLIQNGFKMDQFPYATEBIYIIMQSCWAFDSKRKPSFPNLTSFLGCQL 960
QY 948 ADAEEMYNQVDRVSECHTYQNRR 973
Db 961 ADAEEMYNQVDRVSECHTYQNRR 986

RESULT 15

AAW63589
ID AAW63589 standard; Protein; 986 AA.
XX AC AAW63589;
XX DT 12-OCT-1998 (first entry)
XX DE Human receptor type protein kinase FLT3 protein SEQ ID NO:20.
XX KW Human; receptor type protein kinase; FLT3; diagnosis; leukaemia;
XX KW membrane-vicinal region; tyrosine kinase; juxtamembrane region.
XX OS Homo sapiens.
XX PN WO9817808-A1.
XX PD 30-APR-1998.
XX PF 13-OCT-1997; 97MO-JP03667.
XX PR 18-OCT-1996; 96JP-0297329.
XX PA (TAKI) TAKARA SHUZO CO LTD.
XX PI Yokota S;
XX DR WPI; 1998-362333/31.
XX DR N-PSDB; AAV39042.

XX Nucleic acid sequences encoding receptor type protein kinase -
XX useful for diagnosing adult acute leukaemia of, e.g. FAB sub-classes
XX M2
XX PS Claim 6; Page 50-55; 80pp; Japanese.
XX CC New nucleic acid sequences have been isolated which encode receptor type
XX CC protein kinases (especially a tyrosine kinases) having tandem repeats in
XX CC the juxtamembrane region. Also described in the present invention are:
XX CC (1) any one of 10 amino acid (aa) sequences (AAW63580 to AAW63589)
XX CC representing FLT3 juxtamembrane receptor type protein kinases found in

CC leukaemia patients of FAB subclasses M2, M4 or M5 (nucleic acid sequences
CC encoding the proteins are given in AAV39028 to AAV39042); (2) antibodies
CC recognising the kinases or their portions including the tandem repeat
CC region; (3) nucleic acid sequences hybridising with the nucleic acids;
CC (4) a method for detecting the nucleic acid sequences in human tissue
CC samples by: (i) extracting nucleic acid; (ii) amplifying the kinase
CC gene, and (iii) determining the size of the gene for comparing the size
CC of the normal gene not containing tandem repeats, and (5) kits for
CC carrying out the detection. The products and methods may be used for
CC diagnosing adult acute leukaemia of FAB subclasses M2, M4 and M5.
XX

QY Sequence 986 AA;

Query Match 97.8%; Score 5157.5; DB 19; Length 986;
Best Local Similarity 98.7%; Pred. No. 0;
Matches 973; Conservative 0; Mismatches 0; Indels 13; Gaps 1;

QY 1 MPALARDAGTVPLLVVFSAMIFGTITNODLPVTKCVLINHKNDSSVSKSSSYPMVSESP 60
Db 1 MPALARDAGTVPLLVVFSAMIFGTITNODLPVTKCVLINHKNDSSVSKSSSYPMVSESP 60
QY 61 EDLGCALRPOSSGTVVEAAAEVDVVSASITLQVLVDAPGNIISCLWVFKHSLNCQHPFDL 120
Db 61 EDLGCALRPOSSGTVVEAAAEVDVVSASITLQVLVDAPGNIISCLWVFKHSLNCQHPFDL 120
QY 121 QNRGVSMVLKMTETQAGEYLLFIOSEATNYTLFTVSIRNTLLYTLRRPYFRKMNQD 180
Db 121 QNRGVSMVLKMTETQAGEYLLFIOSEATNYTLFTVSIRNTLLYTLRRPYFRKMNQD 180
QY 181 ALVCISESPEPIVEWVLCDSQESCKEESPAVVKKEKVLHELFGTDIRCCARNELGRE 240
Db 181 ALVCISESPEPIVEWVLCDSQESCKEESPAVVKKEKVLHELFGTDIRCCARNELGRE 240
QY 241 CTRLFITIDLNOTPOTTLPOLFLKVGEPMLIRCAVHVNHGFGLTWELENKALEEGNFEM 300
Db 241 CTRLFITIDLNOTPOTTLPOLFLKVGEPMLIRCAVHVNHGFGLTWELENKALEEGNFEM 300
QY 301 STYSTNRTMIRILFAFVSSVARNDTGYTCSSSKHPSQSALVTIVGKGFNATNSSDEYE 360
Db 301 STYSTNRTMIRILFAFVSSVARNDTGYTCSSSKHPSQSALVTIVGKGFNATNSSDEYE 360
QY 361 IQDYEEFCFSVRKAYPOIRCTWTFSRKSPCPQKGLDNGYSISKFCNHHQGPGEYIFHA 420
Db 361 IQDYEEFCFSVRKAYPOIRCTWTFSRKSPCPQKGLDNGYSISKFCNHHQGPGEYIFHA 420
QY 421 ENDDAQFTKMTLNIIRKQVLAESAASQASCFSDGYPSPSWTKKCDKSPNCTEITE 480
Db 421 ENDDAQFTKMTLNIIRKQVLAESAASQASCFSDGYPSPSWTKKCDKSPNCTEITE 480
QY 481 GVMNRKANRKFQGVSSSTLNMSEAIKGLVKCCAYNSLGTSCETILLNSPGFPFPFIQD 540
Db 481 GVMNRKANRKFQGVSSSTLNMSEAIKGLVKCCAYNSLGTSCETILLNSPGFPFPFIQD 540
QY 541 NISFYATIGVCLLFIVVLTLLCHYKQFRYESQLQWVQVGTSSDNEYFYVDREYEYD 587
Db 541 NISFYATIGVCLLFIVVLTLLCHYKQFRYESQLQWVQVGTSSDNEYFYVDREYEYD 600
QY 588 EYFYVDREYEYDLKWEFFRENLEFGVLSGAFGKVMNATAYGISTGYSIQVAVKMLK 647
Db 601 EYFYVDREYEYDLKWEFFRENLEFGVLSGAFGKVMNATAYGISTGYSIQVAVKMLK 660
QY 648 EKADSSEREALMSLKMWTQLGSHENIVNLLGACTLSGPIYLIPEYCCYGDLLNLYRSKR 707
Db 661 EKADSSEREALMSLKMWTQLGSHENIVNLLGACTLSGPIYLIPEYCCYGDLLNLYRSKR 720
QY 708 EKPHRTWTEIFKEHNFSFYPTFQSHPNSSMPGSRVQIHPDSDQISGLHGNFSHSEDEIE 767
Db 721 EKPHRTWTEIFKEHNFSFYPTFQSHPNSSMPGSRVQIHPDSDQISGLHGNFSHSEDEIE 780
QY 768 YENOKLEEBEDLNVLTFEDLLCFAYQVAKGMFEFLFKSCVHRDLAARNVLVTHGKVVKI 827
Db 781 YENOKLEEBEDLNVLTFEDLLCFAYQVAKGMFEFLFKSCVHRDLAARNVLVTHGKVVKI 840

QY	828	CDFGGLARDIMSDSNVVRGNARLPVKWMAPELSLFEIGIYTIKSDVWSYGILLWEIFSLGVN	887
Db	841	CDFGGLARDIMSDSNVVRGNARLPVKWMAPELSLFEIGIYTIKSDVWSYGILLWEIFSLGVN	900
QY	888	PYPGIPVDANFYKLIQNGFKMDOPFYATEEIIIMOSCWAFDSRKRPSPPNLTSLGCOL	947
Db	901	PYPGIPVDANFYKLIQNGFKMDOPFYATEEIIIMOSCWAFDSRKRPSPPNLTSLGCOL	960
QY	948	ADAEEMYQNVDRVSECPHTYQNR	973
Db	961	ADAEEMYQNVDRVSECPHTYQNR	986

Search completed: August 26, 2003, 07:24:13
Job time : 66.5315 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2003, 07:27:59 ; Search time 45.5229 Seconds
(without alignments)
2878.506 Million cell updates/sec

Title: US-09-919-408A-4
Perfect score: 5274
Sequence: 1 MPALARDAGTVPLLVFSAM.....PFSREMDLGLSPQAQVEDS 993

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 497079 seqs, 131961718 residues

Total number of hits satisfying chosen parameters: 497079

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : Published Applications AA.*
- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
 - 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
 - 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
 - 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
 - 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
 - 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
 - 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
 - 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
 - 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
 - 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
 - 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
 - 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
 - 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
 - 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
 - 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
 - 16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
 - 17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
 - 18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5274	100.0	993	9	US-09-919-408-4
2	5274	100.0	993	10	US-09-872-136-4
3	4429.5	84.0	992	9	US-09-919-408-2
4	4429.5	84.0	992	10	US-09-872-136-2
5	1266	24.0	975	15	US-10-192-867-2
6	1251	23.7	976	14	US-10-099-895-32
7	1251	23.7	976	15	US-10-192-867-4
8	1216.5	23.1	972	10	US-09-944-807-10
9	1166.5	22.1	1088	11	US-09-961-403-4
10	1166	22.1	1089	9	US-09-769-987-2
11	1166	22.1	1089	10	US-09-919-497-90
12	1166	22.1	1089	10	US-09-866-510-2
13	1166	22.1	1089	10	US-09-955-363-36
14	1163	22.1	1089	10	US-09-866-510-10
15	1162	22.0	1089	10	US-09-866-510-4

16	1161	22.0	1089	10	US-09-866-510-8	Sequence 8, Appli
17	1160	22.0	1089	10	US-09-866-510-6	Sequence 6, Appli
18	1086.5	20.6	1090	10	US-09-866-510-14	Sequence 14, Appl
19	1086.5	20.6	1106	10	US-09-866-510-22	Sequence 22, Appl
20	1086.5	20.6	1106	10	US-09-955-363-2	Sequence 2, Appli
21	1082.5	20.5	1106	10	US-09-866-510-16	Sequence 16, Appl
22	1081.5	20.5	1106	10	US-09-866-510-20	Sequence 20, Appl
23	1080.5	20.5	1106	10	US-09-866-510-18	Sequence 18, Appl
24	990	18.8	1338	12	US-10-021-660-84	Sequence 84, Appl
25	990	18.8	1338	15	US-10-059-585-44	Sequence 44, Appl
26	957	18.1	386	9	US-09-939-754-6	Sequence 6, Appli
27	957	18.1	386	10	US-09-939-832-6	Sequence 6, Appli
28	957	18.1	386	10	US-09-939-833-6	Sequence 6, Appli
29	951.5	18.0	1356	15	US-10-022-939-2	Sequence 2, Appli
30	951.5	18.0	1356	15	US-10-100-405A-2	Sequence 2, Appli
31	950.5	18.0	1356	11	US-09-969-037-7	Sequence 7, Appli
32	944.5	17.9	1354	15	US-10-262-538-30	Sequence 30, Appl
33	934	17.7	1367	9	US-09-766-678-2	Sequence 2, Appli
34	923	17.5	1367	9	US-09-919-408-6	Sequence 6, Appli
35	923	17.5	1367	10	US-09-872-136-6	Sequence 6, Appli
36	907.5	17.2	1363	11	US-09-375-248-19	Sequence 19, Appl
37	904.5	17.2	1298	10	US-09-982-610-33	Sequence 33, Appl
38	904.5	17.2	1298	15	US-10-262-538-32	Sequence 32, Appl
39	904.5	17.2	1363	11	US-09-375-248-2	Sequence 2, Appli
40	902.5	17.1	1368	15	US-10-105-901-34	Sequence 34, Appl
41	902	17.1	1362	15	US-10-105-901-33	Sequence 33, Appl
42	880.5	16.7	1363	15	US-10-105-901-32	Sequence 32, Appl
43	859.5	16.3	367	9	US-09-939-754-9	Sequence 9, Appli
44	859.5	16.3	367	10	US-09-939-832-9	Sequence 9, Appli
45	859.5	16.3	367	10	US-09-939-833-9	Sequence 9, Appli

ALIGNMENTS

RESULT 1
US-09-919-408-4
; Sequence 4, Application US/09919408
; Patent No. US20020072077A1
; GENERAL INFORMATION:
; APPLICANT: Lemischka, Ihor R.
; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL
; RECEPTORS AND THEIR LIGANDS
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ImClone Systems Incorporated
; STREET: 180 Varick Street
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10014
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/09/919,408
; APPLICATION DATA:
; FILING DATE: 31-Jul-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/977,451
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US 07/906,397
; FILING DATE: 26-JUN-1992
; APPLICATION NUMBER: US PCT/US92/05401
; FILING DATE: 26-JUN-1992
; APPLICATION NUMBER: TW 81102961
; FILING DATE: 15-APR-1992
; APPLICATION NUMBER: US PCT/US92/02750
; FILING DATE: 02-APR-1992
; APPLICATION NUMBER: US 07/813,593
; FILING DATE: 24-DEC-1991

APPLICATION NUMBER: US 07/793,065
FILING DATE: 15-NOV-1991
APPLICATION NUMBER: US 07/728,913
FILING DATE: 28-JUN-1991
APPLICATION NUMBER: US 07/679,666
FILING DATE: 02-APR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Felt, Irving N.
REGISTRATION NUMBER: 28,601
REFERENCE/DOCKET NUMBER: LEM-3-7P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-645-1405
TELEFAX: 212-645-2054
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 993 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-919-408-4

Query Match 100.0%; Score 5274; DB 9; Length 993;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 993; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MPALARDAGTVELLVVFSAMIFGTITNQDLPVVKVLIHKNQNDSSVGKSSYPMVSESP 60
Db 1 MPALARDAGTVELLVVFSAMIFGTITNQDLPVVKVLIHKNQNDSSVGKSSYPMVSESP 60

Qy 61 EDLGCALRPQSSGTGYEAAAEVDVSASITLQVLDAPGNISCLVWFKHSSLNCPHFPL 120
Db 61 EDLGCALRPQSSGTGYEAAAEVDVSASITLQVLDAPGNISCLVWFKHSSLNCPHFPL 120

Qy 121 QNRGVVSVILKMTQAGEYLLFTQSEATNTTLFTVSIKNTLLYLRPFRKMNOD 180
Db 121 QNRGVVSVILKMTQAGEYLLFTQSEATNTTLFTVSIKNTLLYLRPFRKMNOD 180

Qy 181 ALVCISESPVPIVWLVCDQSGESKESPAVVKKEKVLHFLGTDIRCCARNELGRE 240
Db 181 ALVCISESPVPIVWLVCDQSGESKESPAVVKKEKVLHFLGTDIRCCARNELGRE 240

Qy 241 CTRLTIDLNQPTTLPOLFLKVEPLWIRKAVHNVHGFGLTWELNKALEEGNYFEM 300
Db 241 CTRLTIDLNQPTTLPOLFLKVEPLWIRKAVHNVHGFGLTWELNKALEEGNYFEM 300

Qy 301 STYSTNRMTIRILFAPVSSVARNDTGYTCSKSPQSALVTIVGKFINATNSSEYD 360
Db 301 STYSTNRMTIRILFAPVSSVARNDTGYTCSKSPQSALVTIVGKFINATNSSEYD 360

Qy 361 IDQYEEFCFVRFKAYPQIRCTWTFSRKSFPCEQKGLDNGYSISKFCNHKQFGEYIPHA 420
Db 361 IDQYEEFCFVRFKAYPQIRCTWTFSRKSFPCEQKGLDNGYSISKFCNHKQFGEYIPHA 420

Qy 421 ENDDAQFTQMTINIRKQVLAESAASQAFSDGYPLPSWTWKCKSPNCTBEITE 480
Db 421 ENDDAQFTQMTINIRKQVLAESAASQAFSDGYPLPSWTWKCKSPNCTBEITE 480

Qy 481 GVMNRKANRKFVGMVSSSTLNMSAIGFLVKCCAYNSLGTSCETILLNSGPPRFID 540
Db 481 GVMNRKANRKFVGMVSSSTLNMSAIGFLVKCCAYNSLGTSCETILLNSGPPRFID 540

Qy 541 NISFYATIGVCLLFIIVLTLLCHKYKQFRYESQLQVQVGTSSDNEYFYVDREYEYD 600
Db 541 NISFYATIGVCLLFIIVLTLLCHKYKQFRYESQLQVQVGTSSDNEYFYVDREYEYD 600

Qy 601 LKWEFPRENLEFGKVLGSGAFGKVMATAYIGTGVSTQAVVMUKKADSSERIALMS 660
Db 601 LKWEFPRENLEFGKVLGSGAFGKVMATAYIGTGVSTQAVVMUKKADSSERIALMS 660

Qy 661 ELKMTQOLGSHENIVNLLGACTLSGPIYILIFECCYGGDLLNLYRSKREKHTWTIFKE 720
Db 661 ELKMTQOLGSHENIVNLLGACTLSGPIYILIFECCYGGDLLNLYRSKREKHTWTIFKE 720

Qy 721 HNFSEYPTFQSHPNSSMPGSRVQIHPDSQDQISGLHGNFSHSEDEIEYENQKLEEEEDL 780
Db 721 HNFSEYPTFQSHPNSSMPGSRVQIHPDSQDQISGLHGNFSHSEDEIEYENQKLEEEEDL 780

Qy 781 NVLTPEDELLCPAYQVAKGMEFLPKSCVHRDLAARNVLVTHGKVKVVICDFGLARDIMSDS 840
Db 781 NVLTPEDELLCPAYQVAKGMEFLPKSCVHRDLAARNVLVTHGKVKVVICDFGLARDIMSDS 840

Qy 841 NYVVRGNARLPVKWMAPESLFEGYITIKSDVMSYGILLWEIFSLGVNVPYGPVDPANFYK 900
Db 841 NYVVRGNARLPVKWMAPESLFEGYITIKSDVMSYGILLWEIFSLGVNVPYGPVDPANFYK 900

Qy 901 LIQNGFKMDQPPYATEEYIIMQSCWAFDSRKRPSFPNLTSLGCOLADAEEMYNQVNDG 960
Db 901 LIQNGFKMDQPPYATEEYIIMQSCWAFDSRKRPSFPNLTSLGCOLADAEEMYNQVNDG 960

Qy 961 RVSECPHYQNRPPSREMDLGLLSPOAQVEDS 993
Db 961 RVSECPHYQNRPPSREMDLGLLSPOAQVEDS 993

RESULT 2
US-09-872-136-4
; Sequence 4, Application US/09872136
; Patent No. US20020119545A1
; GENERAL INFORMATION:
; APPLICANT: Lemischka, Ihor R.
; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL
; RECEPTORS AND THEIR LIGANDS
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESS: ImClone Systems Incorporated
; STREET: 180 Varick Street
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10014
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/872,136
; FILING DATE: 01-Jun-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/208,786
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US/09/021,324
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US/07/977,451
; FILING DATE: 1992-11-19
; APPLICATION NUMBER: US 07/906,397
; FILING DATE: 26-JUN-1992
; APPLICATION NUMBER: US PCT/US92/05401
; FILING DATE: 26-JUN-1992
; APPLICATION NUMBER: TW 81102961
; FILING DATE: 15-APR-1992
; APPLICATION NUMBER: US PCT/US92/02750
; FILING DATE: 02-APR-1992
; APPLICATION NUMBER: US 07/813,593
; FILING DATE: 24-DEC-1991
; APPLICATION NUMBER: US 07/793,065
; FILING DATE: 15-NOV-1991
; APPLICATION NUMBER: US 07/728,913
; FILING DATE: 28-JUN-1991
; APPLICATION NUMBER: US 07/679,666
; FILING DATE: 02-APR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Felt, Irving N.
; REGISTRATION NUMBER: 28,601

REFERENCE/DOCKET NUMBER: LEM-3-7P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-645-1405
TELEFAX: 212-645-2054
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 993 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-872-136-4

Query Match 100.0%; Score 5274; DB 10; Length 993;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 993; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPALARDAGTVPVLLVFSAMIFGTITNQDLPVTKVCLINHKNDSSVSKSSSYPMVSESP 60
DB 1 MPALARDAGTVPVLLVFSAMIFGTITNQDLPVTKVCLINHKNDSSVSKSSSYPMVSESP 60
QY 61 EDLGCALRQSSGTVEAAAVDVASITLQVLVDAPGNISCLWVFKHSLNCQPHFDL 120
DB 61 EDLGCALRQSSGTVEAAAVDVASITLQVLVDAPGNISCLWVFKHSLNCQPHFDL 120
QY 121 QNRGVVSWILKMTETQAGEYLLFIOSEATNYTLFTVSRNTLLYTLRPPYRKMENQD 180
DB 121 QNRGVVSWILKMTETQAGEYLLFIOSEATNYTLFTVSRNTLLYTLRPPYRKMENQD 180
QY 181 ALVCISESVPEIVVWLCDQGESCKEESPAVVKKEKVLHFLFGTDIRCCARNELGRE 240
DB 181 ALVCISESVPEIVVWLCDQGESCKEESPAVVKKEKVLHFLFGTDIRCCARNELGRE 240
QY 241 CTRLFTIDNQPTTLPOLFLKVBPLWIRKAVHNVHGFGLTWELNKALBEGNYFEM 300
DB 241 CTRLFTIDNQPTTLPOLFLKVBPLWIRKAVHNVHGFGLTWELNKALBEGNYFEM 300
QY 301 STYSTNRTMIRILFAFVSSVARNDTGYTSCSKHPSQSALVTIVGKGFNATNSSEDE 360
DB 301 STYSTNRTMIRILFAFVSSVARNDTGYTSCSKHPSQSALVTIVGKGFNATNSSEDE 360
QY 361 IDQBEFCFSVRFKAYPOIRCTWTFGRKPPCQKGLDNGYSISKFNHKGHGEYIFHA 420
DB 361 IDQBEFCFSVRFKAYPOIRCTWTFGRKPPCQKGLDNGYSISKFNHKGHGEYIFHA 420
QY 421 ENDDAQFTKMTLNIRKQVLAESAQSCFSDGYPLPSWTWKCKSDKSPNCTEITE 480
DB 421 ENDDAQFTKMTLNIRKQVLAESAQSCFSDGYPLPSWTWKCKSDKSPNCTEITE 480
QY 481 GWNRRKANRKFQGVSSSTLNMSEAIKGLVKCCAYNSLGTSCETILLNSPGPPFIQD 540
DB 481 GWNRRKANRKFQGVSSSTLNMSEAIKGLVKCCAYNSLGTSCETILLNSPGPPFIQD 540
QY 541 NISFVATIGVCLLFIIVLLTLLICHYKQFVRESQLOVQVGTSSDNEFYVDREYED 600
DB 541 NISFVATIGVCLLFIIVLLTLLICHYKQFVRESQLOVQVGTSSDNEFYVDREYED 600
QY 601 LKWEFPRENLEFGKVLGSAFGKVMNATAYGISTGVSIGQAVKMLKEXADSSEREA 660
DB 601 LKWEFPRENLEFGKVLGSAFGKVMNATAYGISTGVSIGQAVKMLKEXADSSEREA 660
QY 661 ELKMTQLGSHENIVNLLGACTLSGPIYLIPEYCCYGDLLNLYRSRKRPHRTWTIFKE 720
DB 661 ELKMTQLGSHENIVNLLGACTLSGPIYLIPEYCCYGDLLNLYRSRKRPHRTWTIFKE 720
QY 721 HNFSPYPTQSPNSMPGSRVQHPDSDQSLGHLGNSFHSDELEYENQKLEEEEDL 780
DB 721 HNFSPYPTQSPNSMPGSRVQHPDSDQSLGHLGNSFHSDELEYENQKLEEEEDL 780
QY 781 NVLTFTEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLTGHGVKVICDFGLARDIMS 840
DB 781 NVLTFTEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLTGHGVKVICDFGLARDIMS 840

QY 841 NYVVRGNARLPVKWMAPESLFEGIYTIKSDVWSYGILLWEIFSLGVNYPGIPVDANFYK 900
DB 841 NYVVRGNARLPVKWMAPESLFEGIYTIKSDVWSYGILLWEIFSLGVNYPGIPVDANFYK 900
QY 901 LIQNGFKMDQFFYATEBIYIIMQSCWAFDSKRKPSFNLTSFLGCLADAEAMYNVDG 960
DB 901 LIQNGFKMDQFFYATEBIYIIMQSCWAFDSKRKPSFNLTSFLGCLADAEAMYNVDG 960
QY 961 RVSECPHTYQNRPPFSREMDLGLLSPQAQVEDS 993
DB 961 RVSECPHTYQNRPPFSREMDLGLLSPQAQVEDS 993

RESULT 3
US-09-919-408-2
; Sequence 2, Application US/09919408
; Patent No. US2002007207A1
; GENERAL INFORMATION:
; APPLICANT: Lemischka, Ihor R.
; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL
; RECEPTORS AND THEIR LIGANDS
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ImClone Systems Incorporated
; STREET: 180 Varick Street
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10014
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/919,408
; FILING DATE: 31-Jul-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/977,451
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US 07/906,397
; FILING DATE: 26-JUN-1992
; APPLICATION NUMBER: US PCT/US92/05401
; FILING DATE: 26-JUN-1992
; APPLICATION NUMBER: TW 81102961
; FILING DATE: 15-APR-1992
; APPLICATION NUMBER: US PCT/US92/02750
; FILING DATE: 02-APR-1992
; APPLICATION NUMBER: US 07/813,593
; FILING DATE: 24-DEC-1991
; APPLICATION NUMBER: US 07/793,065
; FILING DATE: 15-NOV-1991
; APPLICATION NUMBER: US 07/728,913
; FILING DATE: 28-JUN-1991
; APPLICATION NUMBER: US 07/679,666
; FILING DATE: 02-APR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Feit, Irving N.
; REGISTRATION NUMBER: 28,601
; REFERENCE/DOCKET NUMBER: LEM-3-7P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-645-1405
; TELEFAX: 212-645-2054
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 992 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-919-408-2

Query Match		84.0%;	Score 4429.5;	DB 9;	Length 992;
Best Local Similarity		84.1%;	Pred. No. 0;		
Matches 836;		Conservative 57;	Mismatches 90;	Indels 11;	Gaps 4;
QY	1	MPALA-RDAGTVPLLVVFSAMIFGTTITNODLPVKVLINHKNDSSVGKSSSYPMVSES	59		
DB	1	MPALAQRSDRRLLLVLSWILETVTNODLPVKVLISHENNNGSAGKPSYRMVGRS	60		
QY	60	PEDLGCALRPQSSGTYVEAAAEVVDVASITLQVLVDAPGNISCLVFKHSSLNCPHF	119		
DB	61	PEDLQCTPRQSEGTVEAATVEAESSITLQVLATPGDLSCLVFKHSSLCQCPHF	120		
QY	120	LQNRGVSVVILKMTQAGEYLLFTQSEATNTILFTYSIRNTLLTYLRPFKWMNQ	179		
DB	121	LQNRGIVSMALLNVTOAGEYLLHTQSERANYTLFTVNRDTQYVLRPFKWMNQ	180		
QY	180	DALVCISESPRPIVEMVLCDSQGESCKEESPAVKKEKVLHELFGTDIRCCARNELGR	239		
DB	181	DALLCISEGVPETVEWVLCSSHRECKEBGPVAVRKEKVLHELFGTDIRCCARNALGR	240		
QY	240	ECTRLFTIDLNOTPQTLPLQFLKVGEPWIRCKAVVNHGFLGTWELENKALEEGNYFE	299		
DB	241	ECTKLTIDLNOAPOSTLPQFLKVGEPWIRCKAIVNHGFLGTWELEDKALEEGSYFE	300		
QY	300	MSTYSTNRTMIRILPAFVSSVARNDTGYTCSSSKHPSQSALVTIVKGRINATNSSEY	359		
DB	301	MSTYSTNRTMIRILLAFVSSVGRNDGYTCSSSKHPSQSALVTILEKGFINATSSQEEY	360		
QY	360	EIDVVEFCFVRPKAYPQIRCTWTSTRKSPFCBCEOKGLDNGYSISKFNHKKHOPGEVPH	419		
DB	361	EIDPYEFCFVRPKAYPRIRCTWTISQASFPCEQRLGDEGYSISKPCDHNKAPGEYIFY	420		
QY	420	AENDDAQFTKMTFLNIRRKPOVLAEASASQASCFSDGYPLPSMTWKCKSDKSPNCTEEIT	479		
DB	421	AENDDAQFTKMTFLNIRKPOVLANASASQASCSDDGYPLPSMTWKCKSDKSPNCTEEIP	480		
QY	480	EGVWNRKANRVQGVQWSSSTLMSBAIKGLVKCCAYNSLGTSCETILNSGPPFFIQ	539		
DB	481	EGVWNRKANRVQGVQWSSSTLMSBAIKGLVKCCAYNSMGTSCTIFLNSGPPFFIQ	540		
QY	540	DNISFYATIGVCLLFIIVLTLCHYKKOPFYESOLOQVQVSSDNEYFYVDREYEV	599		
DB	541	DNISFYATIGLCPFIIVLTLCHYKKOPFYESOLOQVQVSSDNEYFYVDREYEV	600		
QY	600	DLKWEPPRENLEFGKVLGSGAFGKVMNATAYGSKTGVSIQVAVKMLKEXADSSERBALM	659		
DB	601	DLKWEPPRENLEFGKVLGSGAFGVMNATAYGSKTGVSIQVAVKMLKEXADSCKEALM	660		
QY	660	SELKMMTQLGSHENIVNLLGACTLSGPIYLIIFYCCYGDLLNLYRSKREKPHRTWTIFPK	719		
DB	661	SELKMMTHLGHNDIVNLLGACTLSGPIYLIIFYCCYGDLLNLYRSKREKPHRTWTIFPK	720		
QY	720	EHNFSFYPTFQSHNSMPSGSRVQIHPDSDQISGLHNSFHSDELEYENOKRL--EEE	777		
DB	721	EHNFSYPTFQSHNSMPSGSRVQLHPDSDQISGLHNSFHSDELEYENOKRLAEEEE	780		
QY	778	EDNLVLTFFEDLLCFAYQVAKGMFLEPKSCVHRDLAARNVLTGHKVKVVICDFGLARDIM	837		
DB	781	EDNLVLTFFEDLLCFAYQVAKGMFLEPKSCVHRDLAARNVLTGHKVKVVICDFGLARDIL	840		
QY	838	SDSNVYVRGNARLPVKWMAPESEFEGYITTKSDVMSYGILLWEIFSGVNPYGPVVDAN	897		
DB	841	SDSNVYVRGNARLPVKWMAPESEFEGYITTKSDVMSYGILLWEIFSGVNPYGPVVDAN	900		
QY	898	FYKLIQNGFQWQBPYATEIYIIMQSCWAFDSKRPSPNLTSFLGCOLADAEAMQYN	957		
DB	901	FYKLIQSGFQWQBPYATEIYIFVQSCWAFDSKRPSPNLTSFLGCOLAEAEAC---	957		
QY	958	VDGRVSECPHTYQNRPRPFREMDLGLLSPOAQVE	991		
DB	958	-----IRTSIHLPKQAAPQQRG-GLRAQSPQRQVK	986		

RESULT 4
US-09-872-136-2
; Sequence 2, Application US/09872136
; Patent No. US20020119545A1
; GENERAL INFORMATION:
; APPLICANT: Lemischka, Ibor R.
; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL
; RECEPTORS AND THEIR LIGANDS
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ImClone Systems Incorporated
; STREET: 180 Varick Street
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10014
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/872,136
; FILING DATE: 01-Jun-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/208,786
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US/09/021,324
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US/07/977,451
; FILING DATE: 1992-11-19
; APPLICATION NUMBER: US 07/906,397
; FILING DATE: 26-JUN-1992
; APPLICATION NUMBER: US PCT/US92/05401
; FILING DATE: 26-JUN-1992
; APPLICATION NUMBER: TW 81102961
; FILING DATE: 15-APR-1992
; APPLICATION NUMBER: US PCT/US92/02750
; FILING DATE: 02-APR-1992
; APPLICATION NUMBER: US 07/813,593
; FILING DATE: 24-DEC-1991
; APPLICATION NUMBER: US 07/793,065
; FILING DATE: 15-NOV-1991
; APPLICATION NUMBER: US 07/728,913
; FILING DATE: 28-JUN-1991
; APPLICATION NUMBER: US 07/679,666
; FILING DATE: 02-APR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Felt, Irving N.
; REGISTRATION NUMBER: 28,601
; REFERENCE/DOCKET NUMBER: LEM-3-7P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-645-1405
; TELEFAX: 212-645-2054
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 992 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-872-136-2

Query Match		84.0%;	Score 4429.5;	DB 10;	Length 992;
Best Local Similarity		84.1%;	Pred. No. 0;		
Matches 836;		Conservative 57;	Mismatches 90;	Indels 11;	Gaps 4;
QY	1	MPALA-RDAGTVPLLVVFSAMIFGTTITNODLPVKVLINHKNDSSVGKSSSYPMVSES	59		
DB	1	MPALAQRSDRRLLLVLSWILETVTNODLPVKVLISHENNNGSAGKPSYRMVGRS	60		
QY	60	PEDLGCALRPQSSGTYVEAAAEVVDVASITLQVLVDAPGNISCLVFKHSSLNCPHF	119		

```

61 PEDLOCTPRQSGEYVYEAATVEAEGSITLQVQLATPCDLSCLVFXHSSLGQCPHD 120
QY LQNRGVSNVILKMTTQAGEYLLFIQSEATNYTLFTVSIINTLYTLRRPYFRMENO 179
Db LQNRGIVSMAILNVTQAGEYLLHIQSEBRANYTLFTVNRVDTQLVXLRPYFRMENO 180
QY DALVCISSEVPEPIVWVLCDSQGSCKEESPAVVKKEKVLHELFGTDIRCCARNEGR 239
Db DALLCISGEVPEPTVWVLCSSHRESCKEESPAVVRKEKVLHELFGTDIRCCARNALGR 240
QY ECTRLFTIDLNTOTPTTLPOLFLKVGEPLEIRCKAVHVNHGFLTWELNKALEEGNYFE 299
Db ECTKLTIDLNOAPOSTLTPOLFLKVGEPLEIRCKALHVNHGFLTWELDKALEEGSYFE 300
QY MSTYSTNRTMIRILFAPVSSVARNDTGYTCSHSPQSALVTIVGKGFINATNSSEY 359
Db MSTYSTNRTMIRILFAPVSSVARNDTGYTCSHSPQSALVTILEKGFINATNSSEY 360
QY BIDQVEBFCFSVRFPKAYPOIRCTWTFSRKSPFCBQKGLDNGYSISKFNHKKHOPGEYIIF 419
Db BIDPYEKEFCFSVRFPKAYPRICTWTFSRKSPFCBQKGLDNGYSISKFDHKNKPGYIIFY 420
QY AENDDAQFTKMTLIRRRKQVLAESAQSCFSDGYPLPSMTWKKCSKSPNCTEET 479
Db AENDDAQFTKMTLIRRRKQVLAESAQSCFSDGYPLPSMTWKKCSKSPNCTEET 480
QY EGVNRRKANRKFVGGWSSSTLNMSAEIKGFLVKCCAYNSLGTSCETILLNSPGPPFFIQ 539
Db EGVNRRKANRKFVGGWSSSTLNMSAEIKGFLVKCCAYNSLGTSCETILLNSPGPPFFIQ 540
QY DNISFYATIGVCLLFTIVLTLCHIKYKQFYESOLQVQVGTSSDNFYFYVDREY 599
Db DNISFYATIGVCLLFTIVLTLCHIKYKQFYESOLQVQVGTSSDNFYFYVDREY 600
QY DLKWFPPRENLEFGKVLGSGAGKVMNATAYIGSKTGVSIQAVKMLKEKADSSEREA 659
Db DLKWFPPRENLEFGKVLGSGAGKVMNATAYIGSKTGVSIQAVKMLKEKADSSEREA 660
QY SELKMTOLGSHENTVNLGACTLSGPVLIPEYCCYGGDLLNLYLRKREKFRHTWTEIFK 719
Db SELKMTOLGSHENTVNLGACTLSGPVLIPEYCCYGGDLLNLYLRKREKFRHTWTEIFK 720
QY EHNFFYPTFQSHPNSSMPGSRVQIHPDSQIISGLHNSFHSEDEIYENOKRL--EEE 777
Db EHNFFYPTFQSHPNSSMPGSRVQIHPDSQIISGLHNSFHSEDEIYENOKRLAE 780
QY EDNLVLTLEDLLCFAYQVAKGHEFLFKSCVHRDLAARNVLTGHKVKIKCDFGLARDIM 837
Db EDNLVLTLEDLLCFAYQVAKGHEFLFKSCVHRDLAARNVLTGHKVKIKCDFGLARDIL 840
QY SDSNYYVRGNARLPVKWMAPESLFEGIYTIKSDVMSYGILLWEIPLSGVNPYGPVVDAN 897
Db SDSNYYVRGNARLPVKWMAPESLFEGIYTIKSDVMSYGILLWEIPLSGVNPYGPVVDAN 900
QY FYKLQNGFKMDQPPVATBEIYIIMOSWAFPSRKRPSFPNLTSLFGLQADAEEMAYQN 957
Db FYKLQNGFKMDQPPVATBEIYIIMOSWAFPSRKRPSFPNLTSLFGLQADAEEMAYQN 957
QY VDRGVSECHTYONRRPFRSREMDLGLLSPOAQVE 991
Db VDRGVSECHTYONRRPFRSREMDLGLLSPOAQVE 991

```

RESULT 5

```

US-10-192-867-2
; Sequence 2, Application US/10192867
; Publication No. US20030084466A1
; GENERAL INFORMATION:
; APPLICANT: BLUME-JENSEN, Peter
; APPLICANT: HUNTER, Tony
; TITLE OF INVENTION: SYSTEM AND METHOD FOR CONTROLLING MALE FERTILITY
; FILE REFERENCE: SALKINS.002C1

```

```

; CURRENT APPLICATION NUMBER: US/10/192,867
; CURRENT FILING DATE: 2002-07-08
; PRIOR APPLICATION NUMBER: US 60/175,625
; 2000-01-11
; PRIOR APPLICATION NUMBER: PCT/US01/00573
; 2001-01-08
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 975
; TYPE: PRT
; ORGANISM: MUS MUSCULUS
US-10-192-867-2

Query Match      24.0%; Score 1266; DB 15; Length 975;
Best Local Similarity 33.2%; Pred. No. 2.8e-94;
Matches 336; Conservative 173; Mismatches 330; Indels 172; Gaps 41;

QY 35 CVLINHKNNDSSVGKSSSYPMVSESPEDLGALRPOSSGTVVEAAAVEVDVDSASITLQVL 94
Db 12 CVLLVLLRGQTATSPASBPSPPP-----SIHPAQ-----ELIVEAGDTL--- 54
QY 95 VDAPGNISCL-----WPKHSSLNCQPHFDLQNRGVSNMVLKMTTQAGEYLLFIQSE 148
Db 55 -----SLTICIDPDFVRWTFK-----TYFN-----EMVENKKNWIIQ-EKAE 89
QY 149 ATNYTILFTVSIINTL-----LYTLRRPYFRMENQDQDALVCISESVPEPIVE 195
Db 90 ATR-TGTVTCNSNGLTSSIVYVFRDPAKLFLVGLPLFGK-EDSDALVRCLPTDPO-VSN 146
QY 196 WVLCDQSQESCKEESPAV-----VKKEEVKVLHELFGTDIRCCARNELGRECTRL--- 244
Db 147 YSLIECDGKSLPTDITFVNPKNAGITIKNVKRAYHRLC---VRCAAQ-----RDGTWLHSD 199
QY 245 -FTIDLNOT-----POTTLPOL--FLKVGEPLEIRCKAVHVNHGFLTW-----E 286
Db 200 KTLTKVREAIKAIIPVSVETSHLLKKGDTFVVCITKDVSTSVSNMVLKMNPOQPIHQA 259
QY 287 LENKALEEGNYFEMSTYNTMIRILFAPVSSVARNDTGYTCSHSPQSALVT--- 343
Db 260 VKNHSMHRGDF---NYERQETLT-----ISSARVDDSGVFMVYANNTFGSANVTTLK 309
QY 344 IVKGKFINATN--SEDEYIDQYBECFSVRFPKAYPO-IRCTWTFSRKSPFCBQKGLD--- 398
Db 310 VYKGFINISPVKNTTFTVTDGENVDLVVEYEAYPEKQHYMYMRT--SANKGKDYVK 367
QY 399 -NGYSISKFCNKH-----OPGEYIHAENDDAQFTKMTLIRRRKQVLA--BASAS 448
Db 368 SONKNIRIVNQLRLTRLAGTEGGYITFLVSNSDASASVTFFVNTKPBILTYDRING 427
QY 449 QASCFSDGYPLPSWTWKKCSKSPNCTEBITE-GVWNRKANRKFVGGWSSSTLNMSAEI 507
Db 428 MLQCVABGPEPTIDWYFCTGAERCTTPVSPVDVQVQNVSVSPFGKLVQSSIDSSVFR 487
QY 508 KGLVKCCAYNSLGTSCETILLNSPGPPFP-----IQDN-----ISFYATIGVCLLPIV 557
Db 488 HNGTVECKASNDVGKS--SAFEN-----FAFKEQIOAHTLFTPLLIQFVVAAGANGIIVM 541
QY 558 LTLILCHIKYKQFYESOLQMV-QVTGSSDNFYVDPREYEDLVKWEPRENLEEGKVL 616
Db 542 LT-----YKYLQKPMVEVQWVVEEING---NNYVYIDTQLPYDQHKWEPFRNLSFGKTL 594
QY 617 GSGAFGKVMNATAYIGSKTGVSIQAVKMLKEKADSSEREAALMSLKMTTQGLSHENIVN 676
Db 595 GAGAFGKVMNATAYIGSKTGVSIQAVKMLKEKADSSEREAALMSLKMTTQGLSHENIVN 654
QY 677 LLGACTLSGPVLIPEYCCYGGDLLNLYLRKREKFR-----HRTWTFIFKEHNFSPYFTFQ 730
Db 655 LLGACTVGGFTLVITEYCCYGGDLLNLYLRKREKFRSFIKQBEQAEAAALYKLLHSTEPSCD 714
QY 731 SHPNSSM---PGSRVQIHPDSQIISGLHNSFHSEDEIYENOKRLEEDNLVLTFFED 787
Db 715 S-SNEYMDMKPGVSYV-VPTKTKRRSARIDSY-----IERDVTPAIMDEDEL-ALDLD 766

```

QY 788 LLCFAVQAKGMEFLFKSCVHRDLAARNVLVTHGKVVKICDFGLARDIMSDSNVYVRGN 847
DB 767 LLSFSQVAKAMAFASKNCIHRDLAARNILLTHGRITKICDFGLARDIRDSNVYVWGN 826
QY 848 ARLPVKMAPESLFEGLTYIKSDVWSYGLLWEIFSLGVNYPGICPVDAFYKLIQNGFK 907
DB 827 ARLPVKMAPESLFCVYITPESDWSYGLFWELFSLGSSPYGMPVDSKFKYKMIKEGFR 886
QY 908 MDOPFYATBEIYIMOSCMAFDSRKPSPFNLTSLGCOLADAEAEAMYQNV 958
DB 887 MVSPEHAPAEYMDVMTCDWADPLKRTFKQVQVQLIEKQISDSTKHYSNL 937
RESULT 6
US-10-099-895-32
; Sequence 32, Application US/10099895
; Publication No. US2002017166A1
; GENERAL INFORMATION:
; APPLICANT: BERNDT, Michael C
; APPLICANT: STOMSKI, Frank C
; APPLICANT: LOPEZ, Angel F
; APPLICANT: GUTHRIDGE, Mark A
; TITLE OF INVENTION: A BINDING MOTIF OF A RECEPTOR
; FILE REFERENCE: 3991/0K379
; CURRENT APPLICATION NUMBER: US/10/099,895
; CURRENT FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: PCT/AU00/01118
; PRIOR FILING DATE: 2000-09-15
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 32
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-099-895-32

Query Match 23.7%; Score 1251; DB 14; Length 976;
Best Local Similarity 32.7%; Pred. No. 4.8e-93;
Matches 324; Conservative 166; Mismatches 352; Indels 148; Gaps 34;
QY 47 VGKSSYPMVSESPDLGCLARPOSSGTVYEAFAVEVDVSASITLQVLVDAPGNISCLWV 106
DB 20 VQTGSSQPSVSPG-EPSPPSIHFCKSDLI-----VRVGEIRLLCTDGPVK--WT 67
QY 107 FKHSLLNCQHPFDLQNRGVVSMILKMTETQAGYLLFIQSEATN---YTLFTVSIIRNT 163
DB 68 FE-----ILDETNEKNQEWIT-EKAEATNTGKYTCTNKHGLSNS 106
QY 164 L-----LYTLRRPYFRKMNQDALVCISBSPPEPIVEWVLCDQSGBSCKEE----- 209
DB 107 IYVVRDPKFLVDRSLYK-EDNDTLVRCPLTDPB-VTNYSLKGCQKPLPKDLRFIP 164
QY 210 ---SPAVVKEEKVHLFGTDIRCCARNELGRECTRLFTIDL-----NOTPQTLPQ--L 260
DB 165 DPRAGIMIKSVKRAYHRLC---LHCSVDQEGKSVLSEKFLKVPAPKAVPVVSVSKASY 221
QY 261 FLKVGPELWIRCAVHVNHGFGTLWELN---KALEEGNYFEMSTYSTNRTMIRILFAFV 317
DB 222 LLREGGEFTVCTIKDVSSSVYSTWKRENSQTKLQEKYNSWHHGDFFNYERQAT-----LTI 277
QY 318 SSVARNDDTYTTCSSSKHPQSALVT---IVGKGFINA-----TNSSEYEDIOYE 365
DB 278 SSARVNDSGVPMCVANNTFGSANVTTLVVDVKGFINIFPMINTTVFVNDGENVDL----- 333
QY 366 EPCFSVRFAKYPQ-IRCTWTFSRKSF--PCEQKGLDNGYSISKFNKH-----OPQE 415
DB 334 ---IVEYEAFKPEHQOVIYMNRTFTDKWEDYKSESNIRYVSVLSHLTRLKGTGGT 389
QY 416 YIHAENDDAQFTKMFILNTRRRQVLA--EASASQSCFSDGYPLBSWTWKCSKSPN 473
DB 390 YTLFVNSDVAFAAFNVYNTKFEILTYDRLVNGMLQCVAAGPEPTIDWYFCFGTEQR 449

QY 474 CTEEITE-GVMNRKANRKYFGQWVSSSTLNMSEAIKGFVLKCCAYNSLGTSCETILLNSP 532
DB 450 CSASVLFPVQTLNNSGPPGKLVQSSIDSSAFKNGTVECKAYNDVGT--SAYFN-- 505
QY 533 GPPFPFTODN-----ISFVATIGVCLLFTIVLTLCHYKQKQFYESQLOM 578
DB 506 --PAFKGNKEQIHPHTLFTPLLIGFVIVAGMCIIVMLT-----YKYLQKPMYEVQWKV 559
QY 579 V-QVTSSDNEYFYVDFREYEDLKWEFPRENLEFGKVLGSGAFGKVMNATAYGISKTGV 637
DB 560 VEEING---NNYVYIDPTQLPYDKWEPFRNRLSFGKTLGAGAFGKVEATAYGLIKSDA 616
QY 638 SIQVAVKMLKEKADSSERELAMSELKMTOLGSHENIVNLGACTLSGLPIYLFEYCCYG 697
DB 617 AMTVAVKMLKPSAHLTEREALMSELKVLSYLGNHMINVLGACTTGGPTLVITEYCCYG 676
QY 698 DLLNLYRSKREKPHRTWTE-----IFKEHNPFPYPTFQSHPNSSM---PGSREVOIHPD 748
DB 677 DLLNFLRRKDSFICKQEDHAEALYKLLHSEKSSCDSTNEYMDMKFGVSV-VPTK 735
QY 749 SDQISGLHGNPSHSEDEIYENOKRLEEBEDNLVLTPELDFCFAYQVAKGMEFLFKSCV 808
DB 736 ADKRRSVRIGSY-----IERDVTPIAWEDEL-ALDLEDLSFSYQVAKGMAFLASKNCI 789
QY 809 HRDLAARNVLVTHGKVVKICDFGLARDIMSDSNVYVRGNARLPVKKWAPESLFEIYTIK 868
DB 790 HRDLAARNILLTHGRITKICDFGLARDIKNDSYVVKGNARLPVKKWAPESIFNCVYTFE 849
QY 869 SDVWSYGLLWEIFSLGVNYPGICPVDAFYKLIQNGKMDQFPYATEEYIIMOSQWAF 928
DB 850 SDVWSYGLFWELFSLGSSPYGMPVDSKFKYKMIKEGFRMLSPHAPAEYMDIMKTCDWA 909
QY 929 DSRKRPSFNLTSFLGCOLADAEAEAMYQNV 958
DB 910 DPLKRTFKQIVQLIEKQISESTNHIYSNL 939
RESULT 7
US-10-192-867-4
; Sequence 4, Application US/10192867
; Publication No. US20030084466A1
; GENERAL INFORMATION:
; APPLICANT: BLUME-JENSEN, Peter
; APPLICANT: HUNTER, Tony
; TITLE OF INVENTION: SYSTEM AND METHOD FOR CONTROLLING MALE FERTILITY
; FILE REFERENCE: SALKINS.002C1
; CURRENT APPLICATION NUMBER: US/10/192,867
; CURRENT FILING DATE: 2002-07-08
; PRIOR APPLICATION NUMBER: US 60/175,625
; 2000-01-11
; PRIOR APPLICATION NUMBER: PCT/US01/00573
; 2001-01-08
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 976
; TYPE: PRT
; ORGANISM: HOMO SAPIENS
US-10-192-867-4
Query Match 23.7%; Score 1251; DB 15; Length 976;
Best Local Similarity 32.7%; Pred. No. 4.8e-93;
Matches 324; Conservative 166; Mismatches 352; Indels 148; Gaps 34;
QY 47 VGKSSYPMVSESPDLGCLARPOSSGTVYEAFAVEVDVSASITLQVLVDAPGNISCLWV 106
DB 20 VQTGSSQPSVSPG-EPSPPSIHFCKSDLI-----VRVGEIRLLCTDGPVK--WT 67
QY 107 FKHSLLNCQHPFDLQNRGVVSMILKMTETQAGYLLFIQSEATN---YTLFTVSIIRNT 163
DB 68 FE-----ILDETNEKNQEWIT-EKAEATNTGKYTCTNKHGLSNS 106
QY 164 L-----LYTLRRPYFRKMNQDALVCISBSPPEPIVEWVLCDQSGBSCKEE----- 209

QY 868 KSDVMSYGILLWEIFSLGPNVPGIPVDANFYKLIQNGFKMDOPFYATEEIIYIMOSCA 927
DB 835 QSDVMSYGILLWEIFSLGPNVPGIPVDANFYKLIQNGFKMDOPFYATEEIIYIMOSCA 894
QY 928 FDSRKRPSFNLTSFLGCOL-ADAEAMYQNV 958
DB 895 LEPTHRPTFOQICSFLQBOAQEDRERYTNL 926

RESULT 9
US-09-961-403-4
; Sequence 4, Application US/09961403
; Publication No. US2003007589A1
; GENERAL INFORMATION:
; APPLICANT: HE-STUMPP, HOLGER
; APPLICANT: HAENDLER, BERNARD
; APPLICANT: KRAETZSCHMAR, JOERN
; APPLICANT: KREFT, BERTHOLT
; APPLICANT: WINTERHAGER, ELKE
; APPLICANT: REGIDOR, PEDRO
; APPLICANT: SCOTTI, SIMONE
; TITLE OF INVENTION: METHOD FOR IN VITRO DIAGNOSIS OF ENDOMETRIOSIS
; FILE REFERENCE: SCH-1789
; CURRENT APPLICATION NUMBER: US/09/961,403
; CURRENT FILING DATE: 2001-09-25
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 4
; LENGTH: 1088
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-961-403-4

Query Match 22.1%; Score 1166.5; DB 11; Length 1088;
Best Local Similarity 30.4%; Pred. No. 4.4e-86;
Matches 336; Conservative 165; Mismatches 365; Indels 241; Gaps 36;

QY 12 PLLVVFSAMIFG---TITNODLPVI-----KCVLINHKNDSSVGKSS---SYPMVSES 59
DB 6 PAFLVLGCLLTGLSLILCOLSLPLPNEKVNQVQLNSSFSLRCFGESEVSWQPMSEEE 65
QY 60 PEDLGCALRPQSG---TVYEAANAEDVDVSAITQLVLVDAPGNISCLWPFKHSSINCO 116
DB 66 SSDVEIRNEENSGLFVTVLEVSS-----ASAHT-----GLVTCYNNHTQTEEN--- 110
QY 117 HFDLQNRGV-----VSMWILKMTETQAGGYLLFIQSEATNYTILFTVSRNTLLYTL 168
DB 111 --ELEGRIHYIVPDPDAFVPLGWT-----DYLIVE----- 141
QY 169 RPYFRKMNQDALVCISESVPEPIVWVCLDSQG-----ESCKE 208
DB 142 -----DDSAIIPCRTPDPTPV--TLHNSGVVPASVDSRQGFNGTFTVGPVIC-- 189
QY 209 ESPAVVKEEKVLHFLGTDIRCCARNELGRECTRLFTIDLNQTPOTTLPOFLKVGEP 268
DB 190 EATVKGKFPQTIPFNVYALK-----ATSELDLEMEALKTV-----YKSGETI 231
QY 269 WIRCKAVHNVHFGLTW-----ELENKA-----LEGNYFEMSTYTNTRMIRILPAF-VSSV 320
DB 232 VVTC-AVFNEVVDLQWTFVGEVKGKITMLEE-----IKVPSIKLVYTLTVPEA 280
QY 321 ARNDTGYTCCSSK-----HPQSALVTIVGKGFINATNSSEDEY-IDQYEFPCFSVRPK 374
DB 281 TVKDSGDVECAARQATREVKEMKVVITISVHEKGFIEIKPTFSQLEAVNLHVKHFVVE 340
QY 375 AYPQIRCTWTFERKFPCEQKLDNGYSIS-----KPCNHKIQ-POEYIF 418
DB 341 AYPPIRISWLKNLTL-----IENLTETDVEKIEIRYSKULIRAKEEDSGHYTI 394
QY 419 HAENDDAQFTKMTNIRKPOVL-----AEASASQASCFSDGYPLPSWTWTKKCDKS 471
DB 395 VAQNEADAVKSYTFELLTQVPSILDVDDHGGSTGGQTVRCAGPLPDIEMWICKD-I 453

QY 472 PNCTEEITEGVNKRKRVQWYSS-----STLNMSEAIKGLYKCCAY 517
DB 454 KCCNNETS---WTLIANN-----VSNITTEIHSRDRSTVEGRVTFKAVEETIAVRCLAK 504
QY 518 NSLGTSCETILLNSPGPPFFIQDNISFVATICVCLLFIVLTLCHYKVKQFRYESQIQ 577
DB 505 NLGRENRELKVA---PTLRESELTVAANVLVIVLIVLIVLIVLIVLIVLIVLIVLIVLIV 560
QY 578 MVQVTGSSDNRYFYVDPREYEDLKWEPFRENLEFGKVLGSGAFGKVMNATAYIGSKTV 637
DB 561 VIESISPDGHEIYYVDPMLPYDSRWEFFRDGLVGLVGLGSGAFGKVTAGYLSRSQP 620
QY 638 SIQAVAKMLKESADSSEREALMSLKMTOLGSHENIVNLLGACTLSGPIYILIFRYCCY 697
DB 621 VMKAVAKMLKPTARSSEKQALMSLKMTOLGSHENIVNLLGACTLSGPIYIITEYCFY 680
QY 698 DLLNLYLRKREKF-----HRTWTEIFKEHNFSFYPTFOSHPN 734
DB 681 DLVNYLHKNRDSFLSHHPEKPKKELDIPLGNADSTRSYVILSPENNGDYMDKQADTT 740
QY 735 SSMPSREVVQIHPDSQISGLHGN--SFHSEDEIEYENQKRLSEEDLNLVTFEDLLCPA 792
DB 741 QYVPMLEKEVSKYSDIQRSLYDRPASVYKCKSMLDSE-VKLLSDNSSEGLTLLDLSFT 799
QY 793 YQVAKGMEFLBPKSCVHRDLAARNVLVTHGKVKIKCDPGLARDIMSDSNYVVRGNARLP 852
DB 800 YQVARGMEFLASKNCVHRDLAARNVLLAQGKTKIKCDPGLARDIMSDSNYVSGSTFLPV 859
QY 853 KMAPESLFEGIYTIKSDVMSYGILLWEIFSLGPNVPGIPVDANFYKLIQNGFKMDOP 912
DB 860 KMAPESLFDMYTLSDVMSYGILLWEIFSLGPNVPGIPVDANFYKLIQNGFKMDOP 919
QY 913 YATEEIIYIMOSCAFPDSRKRPSFNLTSFLGCOLADAEAMYQNV--DGRVSECPH--- 967
DB 920 HATSEVVEIMVKMNSSEKPSFYHLSEIVENLLPGQYKYSYKTHLDFLKSDEHFAVAR 979
QY 968 -----TYQNRPRFSPREMDLGL 983
DB 980 MYDSDNAYIGVTYKNEEDKLDWEGGL 1006

RESULT 10
US-09-769-987-2
; Sequence 2, Application US/09769987
; Patent No. US20020055129A1
; GENERAL INFORMATION:
; APPLICANT: Matsui, Toshimitsu
; APPLICANT: Aaronson, Stuart A.
; APPLICANT: Pierce, Jacalyn H.
; TITLE OF INVENTION: Antibodies for the Alpha Platelet-Derived Growth Factor Receptor;
; FILE REFERENCE: 14014.026602
; CURRENT APPLICATION NUMBER: US/09/769,987
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 08/460,656
; PRIOR FILING DATE: 1995-06-02
; PRIOR APPLICATION NUMBER: US 08/439,095
; PRIOR FILING DATE: 1995-05-11
; PRIOR APPLICATION NUMBER: US 07/915,884
; PRIOR FILING DATE: 1992-07-20
; PRIOR APPLICATION NUMBER: US 07/308,282
; PRIOR FILING DATE: 1989-02-09
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1089
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence; No. US20020055129A1 =
; OTHER INFORMATION: synthetic construct
US-09-769-987-2

Query Match 22.1%; Score 1166; DB 9; Length 1089;

Db 505 NLGAENRELKVA-----PTLRSELTAAAVLVLLVIVLIIIVLVLVWVKQKPRYIEWR 560
 Qy 578 MVQVTGSDNEYFYVDREYEDLKWEPRENLEFGKVLGSGAFKGMNATAYGSKTV 637
 Db 561 VIESIPDGHEIYVDPMLQPYDSRWEFRDGLVLRVLGSGAFKGMNATAYGSKTV 620
 Qy 638 SIQVAVKMLKEKADSSEREAALMSLKMOTLGSHENIIVNLLGACTLSGPIYILFEYCCYG 697
 Db 621 VMKAVKMLKPTARSSEREAALMSLKMOTLGSHENIIVNLLGACTLSGPIYILFEYCCYG 680
 Qy 698 DLLNVLRSKREK-----HRTWTEIFKEHNFSPYPTFQSHPN 734
 Db 681 DLVNLHKNRDSFLSHHPEKPKKELDI FGLNPADESTRSYVLSFENNDDYMDKQADTT 740
 Qy 735 SSMPGSEVQIHPDSDQISGLHGN--SPHSEDEIEYENOKRLEEBEDNLVLFEDLLCFA 792
 Db 741 QYVPMLEKESKYSIDQRSYDRPASYYKKSMLOSE-VKNLLSDNSEGGLTLLDLSFT 799
 Qy 793 YQVAKGMEFLPKSCVHRDLAARNVLVTHGKVVKICDFGLARDIMSDSNYYVGRNARLPV 852
 Db 800 YQVARGMEFLASKKNCVHRDLAARNVLLAQQKIVKICDFGLARDIMHDSNYVSKGSTFLPV 859
 Qy 853 KWMAPESIFEGITYIKSDVWSYGILLWEIFSGVNPYPCIPVDANFYKLIQNGFKMDQPF 912
 Db 860 KWMAPESIFDNLTYTLLSDVWSYGILLWEIFSGVNPYPCIPVDANFYKLIQNGFKMDQPF 919
 Qy 913 YATSEIYIMOSCAFDGRKPSFPNLTSLFCQLADAEAMYNQV--DGRVSECPH--- 967
 Db 920 HATSEVIEIMVKNWSEKREPSFYLSEIVENLLPGQYKYSYKIHLDLFLKSDHPAVAR 979
 Qy 968 -----TYQNRPPFSREMDLGL 983
 Db 980 MRVSDNAYIGVITYKNEEDKLDWEGGL 1007

RESULT 12
 US-09-866-510-2
 ; Sequence 2, Application US/09866510
 ; Patent No. US20020113041
 ; GENERAL INFORMATION:
 ; APPLICANT: KAZLAUSKAS, ANDRIUS
 ; APPLICANT: IKUNO, YASUHI
 ; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING FIBROTIC DISEASES
 ; FILE REFERENCE: ERM-104.01
 ; CURRENT APPLICATION NUMBER: US/09/866,510
 ; PRIORITY FILING DATE: 2001-05-25
 ; PRIOR APPLICATION NUMBER: 60/250,747
 ; PRIOR FILING DATE: 2000-12-01
 ; PRIOR APPLICATION NUMBER: 60/289,103
 ; PRIOR FILING DATE: 2001-05-07
 ; NUMBER OF SEQ ID NOS: 33
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 2
 ; LENGTH: 1089
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-09-866-510-2

Query Match 22.1%; Score 1166; DB 10; Length 1089;
 Best Local Similarity 30.3%; Pred. No. 4.8e-86;
 Matches 336; Conservative 165; Mismatches 365; Indels 242; Gaps 36;
 Qy 12 PLLVVFSAIFG---TITNQLPVI-----KCVLLHNKNDSSVSKSS---SYPMVSES 59
 Db 6 PAFLVGLCLTGLSLICQLSPILPNEKVKVQNLNLSFLRCFGESEVSWQYPMSEEE 65
 Qy 60 PEDLGCALPQSSG---TYVEAAVEVDVSASITLQVLVDAPGNISCLMFKVHSSLNQOP 116
 Db 66 SSDVEIRNEENNSGLPVTLEVSS-----ASAHT-----GLTYCYNHTQTEEN--- 110
 Qy 117 HFDLQNRGV-----VSNVILKMTQAGEVLLFIQSEATNYTLFTVSIANTLLTYL 168
 Db 111 --ELEGRHIYIYVDPDPAFVPLGMT-----DYLIVE----- 141

Qy 169 RRPYFRKMNQDALVCISESVPEPIVWVWVLCDSQ-----ESCKE 208
 Db 142 -----DDSAIIPCTTDPETPV--TLNSEGVPASVDSRQGFNGTFTVGVYIC-- 189
 Qy 209 ESPAVKKEEVLHELFGTDIRCCARNELGRECTRLFTIDLNTQTPTTLPLQLFLKVGEP 268
 Db 190 EATVKGKFKQIIPFNVAALK-----ATSELDLEMEALKTV-----YKSGEIT 231
 Qy 269 WIRKCAVHNHFGGLTW-----ELENKA---LEEGNYFEMSTYSTNRTWIRILFAP-VSSV 320
 Db 232 VVTC-AVFNNEVDLQWTPYGEVKGKGTMLKE-----IKVPSIKLYTTLTVPEA 280
 Qy 321 ARNDGYTTCSSK-----HPSQSNALVTIVGKGFINATNSSEDIY-IDOYEFECFSVRK 374
 Db 281 TVKOSGDYECARQATREVKEKMKVTISVHEKGFIEIKPTFSQLEAVNLHVKHFVVEVR 340
 Qy 375 AVPOICRTWTTSRKSFPCEQKGLDNGYIS-----KFCNHHQK-PGEYIF 418
 Db 341 AYPPPRIISWLNKNTL-----IENLTIITDVEIKQIRYRSKLLIPAKEEDSGHYTI 394
 Qy 419 HAENDDAQFTKMTFLNIRKQVVL-----AEASASQASCFSDGYPLPSWTWKKCSDKS 471
 Db 395 VAQNEADVKSFTFELLTQVPSILLDLDHGGSTGGTVRCTAEGTPLPDIEMWICKD-I 453
 Qy 472 PNCTEITEGVWNRKANRKFQWVSS-----STLNMSEAIKGLVKCCAY 517
 Db 454 KKCNNETS---WTLANN-----VSNITTEIHSRDRSTVEGRVTFKVEETIAVRCLAK 504
 Qy 518 NLSGTSCETILLNSPGPFQIDNLSFYATIGVCLLIVLVLLIICHYKVKOPRYESQ 577
 Db 505 NLLGAENRELKVA-----PTLRSELTAAAVLVLLVIVLIIIVLVLVWVKQKPRYIEWR 560
 Qy 578 MVQVTGSDNEYFYVDREYEDLKWEPRENLEFGKVLGSGAFKGMNATAYGSKTV 637
 Db 561 VIESIPDGHEIYVDPMLQPYDSRWEFRDGLVLRVLGSGAFKGMNATAYGSKTV 620
 Qy 638 SIQVAVKMLKEKADSSEREAALMSLKMOTLGSHENIIVNLLGACTLSGPIYILFEYCCYG 697
 Db 621 VMKAVKMLKPTARSSEREAALMSLKMOTLGSHENIIVNLLGACTLSGPIYILFEYCCYG 680
 Qy 698 DLLNVLRSKREK-----HRTWTEIFKEHNFSPYPTFQSHPN 734
 Db 681 DLVNLHKNRDSFLSHHPEKPKKELDI FGLNPADESTRSYVLSFENNDDYMDKQADTT 740
 Qy 735 SSMPGSEVQIHPDSDQISGLHGN--SPHSEDEIEYENOKRLEEBEDNLVLFEDLLCFA 792
 Db 741 QYVPMLEKESKYSIDQRSYDRPASYYKKSMLOSE-VKNLLSDNSEGGLTLLDLSFT 799
 Qy 793 YQVAKGMEFLPKSCVHRDLAARNVLVTHGKVVKICDFGLARDIMSDSNYYVGRNARLPV 852
 Db 800 YQVARGMEFLASKKNCVHRDLAARNVLLAQQKIVKICDFGLARDIMHDSNYVSKGSTFLPV 859
 Qy 853 KWMAPESIFEGITYIKSDVWSYGILLWEIFSGVNPYPCIPVDANFYKLIQNGFKMDQPF 912
 Db 860 KWMAPESIFDNLTYTLLSDVWSYGILLWEIFSGVNPYPCIPVDANFYKLIQNGFKMDQPF 919
 Qy 913 YATSEIYIMOSCAFDGRKPSFPNLTSLFCQLADAEAMYNQV--DGRVSECPH--- 967
 Db 920 HATSEVIEIMVKNWSEKREPSFYLSEIVENLLPGQYKYSYKIHLDLFLKSDHPAVAR 979
 Qy 968 -----TYQNRPPFSREMDLGL 983
 Db 980 MRVSDNAYIGVITYKNEEDKLDWEGGL 1007

RESULT 13
 US-09-955-363-36
 ; Sequence 36, Application US/09955363
 ; Patent No. US20020173621A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Siedziawski Ph.D., Andrzej 2
 ; Bell, Lillian A.

Kindsvogel Ph.D., Wayne R.
 TITLE OF INVENTION: METHODS OF PRODUCING SECRETED RECEPTOR ANALOGS
 AND BIOLOGICALLY ACTIVE DIMERIZED POLYPEPTIDE
 FUSIONS
 NUMBER OF SEQUENCES: 36
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Seed and Berry
 STREET: 6300 Columbia Center
 CITY: Seattle
 STATE: WA
 COUNTRY: USA
 ZIP: 98104-7092
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent in Release #1.24
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/955,363
 FILING DATE: 18-Sep-2001
 CLASSIFICATION: <Unknown>
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/634,510
 FILING DATE: <Unknown>
 APPLICATION NUMBER: US 07/347,291
 FILING DATE: 02-MAY-1989
 ATTORNEY/AGENT INFORMATION:
 NAME: Maki J.D., David J.
 REGISTRATION NUMBER: 31,392
 REFERENCE/DOCKET NUMBER: 990008.446C3
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 206-622-4900
 TELEFAX: 206-682-6031
 TELEX: 3723836
 INFORMATION FOR SEQ ID NO: 36:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1089 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 SEQUENCE DESCRIPTION: SEQ ID NO: 36:
 US-09-955-363-36

Query Match 22.1%; Score 1166; DB 10; Length 1089;
 Best Local Similarity 30.3%; Pred. No. 4.8e-86;
 Matches 336; Conservative 165; Mismatches 365; Indels 242; Gaps 36;

12 PLLVVFSAIFG---TITNQLPVI-----KCVLNHNKNDSSVCKSS---SYPMVSES 59
 6 PAFLVGLCLTGLSILQLSLPSILPNEKVVQLNSFSLRCPGSESVWQYPMSEEE 65
 60 PEDLGCALRPOSSG---TVYEAAVEVDVSAITLQVLVDAPGNISCLMVFVKHSSINCO 116
 66 SSDVEIRNEENNSGLFVTVLEVSS---ASAHT-----GLYTCYNHNTQTEEN--- 110
 117 HPDLQNGV-----VSMVLKMTQAGYLLFIQGEAANYTILFTVSTRINTLLYTL 168
 111 --ELEGRHIYIYVPDVAFLPGLMT-----DYLVIIE----- 141
 169 RPYFERKMNQDALVCISEVPEPIVWVLCDSQ-----ESCKE 208
 142 -----DDSAIIPCKRTDPTETV--TLHNSEGVVPASVDSROGFNGFTTVGPYIC-- 189
 209 ESPAVVKEEKLHVLFGTDIRCCARNELGRCTRFLTIDLNQTPQTTLPLQFLKVGEP 268
 190 EATVKGKFFQIPFNVAALK-----ATSELDLEWALKTV-----YKSGETI 231
 269 WIRKAVHNVHGFGLTW-----ELENA-----LEEGNYFEMSTYSTNRTMIRILFAP-VSSV 320
 232 VVTC-AVENNEVVDLQWTVPGVKGKITLSE-----IKVPSIKLVYTLTVPEA 280
 321 ARNDTGYTCSSK-----HPSQSLAVTVIGKGFNATNNSSEDEY-IDQYEFECFSVRPK 374

281 TVKSGDYECAARQATREVKEMKQVTISVHEKGFBIKPTFSQLEAVNLHVKHFVVVEVR 340
 375 AYPQIRCTWTFPRSKSPCEQKGLDNGYSIS-----KFCNHHKQ-PGEYIF 418
 341 AYPFPRISLWKNLTL-----IENLTETTDVEKIQEIRYSKLIIRAKEDSGHYTI 394
 419 HAENDDAOFTKMTLNIIRKQOVL-----AEASASQASCFSDGYPLPSWTWKKCSDKS 471
 395 VAQNEADVASYTFELLTQVPSSILDLVDHGHSTGGQTVRCTAEGTPLPDIEWICKD-I 453
 472 PNCTEEITEGVNRRKANRKFQGWSS-----STLNMSEAKGFLVKCCAY 517
 454 KCKNNETS---WTILANN-----VSNIIITEHSRDRSTVEGRVTPAKVEETIAVCLAK 504
 518 NSLGTSCETILLNSGPPPIQDNISFYATIGVCLLFIVVLILLIKHKVKKQPRYSOLQ 577
 505 NLLGAENRELKVA---PTLSELTVAAAVLLVIVLIVLIVLIVLIVLIVLIVLIVLIV 560
 578 MVQVTGSSDNEYFYVDREYEDLKWEPRENLEFGKVLGSGAGKVMNATAYGISKTGV 637
 561 VIESISPDGHEIYYVDPQOLPYDSRWEFPRDGLVLRVLGSGAFKVGSTAYGLSRQP 620
 638 SIQVAVKMLKESKADSSEREAALMSLKMMTQLGSHENIVNLLGACTLSGPIYLIFFEYCCY 697
 621 VMKAVAVKMLKPTARSEKQALMSLKMTHLGPHLNIVNLLGACTSGPIYIITEYCFY 680
 698 DLLNLYLSKREKF-----HRTWTRIFKEHNFSFYPTFQSHPN 734
 681 DLVNYLHKNRDSFLSHHPEKPKKELDIFGLNPADESTRSYVILSPENNGDYMDKQADTT 740
 735 SMPGSRVQIHPDSQISGLHGN--SPHSEDEIEYENOKRLEEEDLNVLPFEDLLCPA 792
 741 QYVPMLEKREKVSYSIDIQRSLYDRPASVYKKSMLDSE-VKNLLSDNNSGLTLLDLLSET 799
 793 YOVAKGMEFLPKSCVHRDLAARNVLTGHGVKVCDFGLARDIMSDSNVVRGNARLPV 852
 800 YQVARGMEFLASKNCVHRDLAARNVLLAQCKVVKICDFGLARDIMSDSNVYKSGTFLPV 859
 853 KMAPESLFEGIYTIKSDVMSYGILLWEIFSLGVNPPYGPIDVNFYKLIQNGFKMDQPF 912
 860 KMAPESIFDNLTYTLLSDVMSYGILLWEIFSLGGTYPYGMVVDSTFYNKIKSGYRMAKPD 919
 913 YATEIYIIMQSWAFDSRKRPSFNLTSLFCQCLADAEEMQNV--DGRVSECPH--- 967
 920 HATSEVSEIMVKMNSPEKRPSTFHLSEIVENLLPGQYKKSVEKIHLDFLKSDHPAVAR 979
 968 -----TYONRRPFSGREMDLGL 983
 980 MRVDSDNAYIGVTYKNEEDKLDWEGGL 1007

RESULT 14
 US-09-866-510-10
 ; Sequence 10, Application US/09866510
 ; Patent No. US20020111304A1
 ; GENERAL INFORMATION:
 ; APPLICANT: KAZLAUSKAS, ANDRIUS
 ; APPLICANT: IKUNO, YASUSHI
 ; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING FIBROTIC DISEASES
 ; FILE REFERENCE: ERM-104.01
 ; CURRENT APPLICATION NUMBER: US/09/866,510
 ; CURRENT FILING DATE: 2001-05-25
 ; PRIOR APPLICATION NUMBER: 60/250,747
 ; PRIOR FILING DATE: 2000-12-01
 ; PRIOR APPLICATION NUMBER: 60/289,103
 ; PRIOR FILING DATE: 2001-05-07
 ; NUMBER OF SEQ ID NOS: 33
 ; SOFTWARE: Patent In Ver. 2.1
 ; SEQ ID NO 10
 ; LENGTH: 1089
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-866-510-10

```
Query Match      22.1%; Score 1163; DB 10; Length 1089;
Best Local Similarity 30.2%; Pred. No. 8.5e-86;
Matches 335; Conservative 166; Mismatches 365; Indels 242; Gaps 36;

QY 12 PLLVFSAMIFG---TITNQDLPVI-----KCVLINHKNDSSVCKSS---SYPMVSES 59
DB 6 PAFLVLGCLLTGLSLILCOLSLPILPNENKVKVQLNSFSRLRCFGESEVSWQYPMSEEE 65

QY 60 PEDLGALRQSSG---TVYEAADVDSASITLQVLVDAPGNISCLWVFKHSSLNCQP 116
DB 66 SSDVEIRNEENSGLPVTVLEVSS-----ASAHT-----GLTCYNNHTQTEEN--- 110

QY 117 HFDLQNRGV-----VSMVILKMTETOAGEYLLFIQSEATNYTILFTVSRINTLLYTL 168
DB 111 --ELEGRHIYIVPDPAFVPLGMT-----DYLIVIVE-----YKSGETI 231

QY 169 RPYFRKMNQDALVCISSEVPEPIVWVLCDSQ-----ESCKE 208
DB 142 -----DDSAIIPCRITTDPTPV--TLHNSGVVPASYSRQGFNGFTTGVGYIC-- 189

QY 209 ESPAVVVKKEKVLHFGTDIRCCARNELGRECTRLFTIDLNOTPTTLPLQFLKVGCEPL 268
DB 190 EATVKGKKEQTIPFNVAULK-----ATSELDLEMEALKTV-----YKSGETI 231

QY 269 WIRKCAVHNVHGFGLTW-----ELENKA-----LEEGNYFEMSTYSTNRTMIRILFAP-VSSV 320
DB 232 VVTC-AVFNEVVLDLQWTPYGEVKGKGITWLEB-----IKVPSIKLVTLTVPEA 280

QY 321 ARNDTGYTCCSSK-----HPSQSALVTIVGKFINATNSSEDEY-IDQVEEFCFSVRFK 374
DB 281 TVKDSGDYECARQAATREVKEMKKTIVSHEKGFIBIKPTFSQLEAVNLHEVKHFVVEVR 340

QY 375 AYPQIRCTWTFSRKSPCEQKGLDNGYSIS-----KFCNHKHQ-PGEYIF 418
DB 341 AYPPPPISLKNLTL-----IENLTETDVEKIQEIYRSKLLIRAKEDSDGHYTI 394

QY 419 HAENDDAQTKMFTLNIIRKPOVL-----AEASASQSCFSDGYPLPSWTWKCKSDKS 471
DB 395 VAQNEADAVKSYTFELLTQVPSILLDLDVDDHGGTGGTQVRCCTAEGTPLDIEWICKD-I 453
QY 472 PNCTEITEGVWNRKANRKFVGQWSS-----STLNMSEAIKGLVKCCAY 517

Query Match      22.0%; Score 1162; DB 10; Length 1089;
Best Local Similarity 30.2%; Pred. No. 1e-85;
Matches 335; Conservative 166; Mismatches 365; Indels 242; Gaps 36;

QY 12 PLLVFSAMIFG---TITNQDLPVI-----KCVLINHKNDSSVCKSS---SYPMVSES 59
DB 6 PAFLVLGCLLTGLSLILCOLSLPILPNENKVKVQLNSFSRLRCFGESEVSWQYPMSEEE 65

QY 60 PEDLGALRQSSG---TVYEAADVDSASITLQVLVDAPGNISCLWVFKHSSLNCQP 116
DB 66 SSDVEIRNEENSGLPVTVLEVSS-----ASAHT-----GLTCYNNHTQTEEN--- 110

QY 117 HFDLQNRGV-----VSMVILKMTETOAGEYLLFIQSEATNYTILFTVSRINTLLYTL 168
DB 111 --ELEGRHIYIVPDPAFVPLGMT-----DYLIVIVE-----YKSGETI 231

QY 169 RPYFRKMNQDALVCISSEVPEPIVWVLCDSQ-----ESCKE 208
DB 142 -----DDSAIIPCRITTDPTPV--TLHNSGVVPASYSRQGFNGFTTGVGYIC-- 189

QY 209 ESPAVVVKKEKVLHFGTDIRCCARNELGRECTRLFTIDLNOTPTTLPLQFLKVGCEPL 268
DB 190 EATVKGKKEQTIPFNVAULK-----ATSELDLEMEALKTV-----YKSGETI 231

QY 269 WIRKCAVHNVHGFGLTW-----ELENKA-----LEEGNYFEMSTYSTNRTMIRILFAP-VSSV 320
DB 232 VVTC-AVFNEVVLDLQWTPYGEVKGKGITWLEB-----IKVPSIKLVTLTVPEA 280

QY 321 ARNDTGYTCCSSK-----HPSQSALVTIVGKFINATNSSEDEY-IDQVEEFCFSVRFK 374
DB 281 TVKDSGDYECARQAATREVKEMKKTIVSHEKGFIBIKPTFSQLEAVNLHEVKHFVVEVR 340

QY 375 AYPQIRCTWTFSRKSPCEQKGLDNGYSIS-----KFCNHKHQ-PGEYIF 418
DB 341 AYPPPPISLKNLTL-----IENLTETDVEKIQEIYRSKLLIRAKEDSDGHYTI 394

QY 419 HAENDDAQTKMFTLNIIRKPOVL-----AEASASQSCFSDGYPLPSWTWKCKSDKS 471
DB 395 VAQNEADAVKSYTFELLTQVPSILLDLDVDDHGGTGGTQVRCCTAEGTPLDIEWICKD-I 453
QY 472 PNCTEITEGVWNRKANRKFVGQWSS-----STLNMSEAIKGLVKCCAY 517
```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2003, 07:21:59 ; Search time 21.5108 Seconds
(without alignments)
1953.187 Million cell updates/sec

Title: US-09-919-408A-4
Perfect score: 5274
Sequence: 1 MPALARDAGTVPLLVFSAM.....PFSREMDLGLSPQAVQVEDS 993

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5274	100.0	993	1 US-07-977-451-4	Sequence 4, Appli
2	5274	100.0	993	1 US-08-252-517-4	Sequence 4, Appli
3	5274	100.0	993	1 US-07-906-397A-4	Sequence 4, Appli
4	5274	100.0	993	1 US-08-601-891-4	Sequence 4, Appli
5	5274	100.0	993	2 US-09-021-324-4	Sequence 4, Appli
6	5274	100.0	993	5 PCT-US92-09893-4	Sequence 4, Appli
7	5274	100.0	1160	5 PCT-US92-05401-4	Sequence 4, Appli
8	5240	99.4	993	1 US-08-222-299-4	Sequence 4, Appli
9	5240	99.4	993	2 US-08-434-878-4	Sequence 4, Appli
10	5240	99.4	993	5 PCT-US95-03718-4	Sequence 4, Appli
11	5238	99.3	993	1 US-08-183-211-2	Sequence 2, Appli
12	5238	99.3	993	5 PCT-US95-00176A-2	Sequence 2, Appli
13	4533.5	86.0	1000	1 US-08-222-299-2	Sequence 2, Appli
14	4533.5	86.0	1000	2 US-08-434-878-2	Sequence 2, Appli
15	4533.5	86.0	1000	5 PCT-US95-03718-2	Sequence 2, Appli
16	4429.5	84.0	992	1 US-07-813-593-2	Sequence 2, Appli
17	4429.5	84.0	992	1 US-07-977-451-2	Sequence 2, Appli
18	4429.5	84.0	992	1 US-07-946-507-2	Sequence 2, Appli
19	4429.5	84.0	992	1 US-08-252-517-2	Sequence 2, Appli
20	4429.5	84.0	992	1 US-07-906-397A-2	Sequence 2, Appli
21	4429.5	84.0	992	1 US-08-601-891-2	Sequence 2, Appli
22	4429.5	84.0	992	2 US-09-021-324-2	Sequence 2, Appli
23	4429.5	84.0	992	5 PCT-US92-02750-2	Sequence 2, Appli
24	4429.5	84.0	992	5 PCT-US92-05401-2	Sequence 2, Appli
25	4429.5	84.0	992	5 PCT-US92-09893-2	Sequence 2, Appli
26	1654	31.4	481	3 US-07-912-122-4	Sequence 4, Appli
27	1654	31.4	481	5 PCT-US93-06404-4	Sequence 4, Appli

28 1251 23.7 976 3 US-08-750-141A-1 Sequence 1, Appli
29 1216.5 23.1 972 3 US-08-750-141A-2 Sequence 2, Appli
30 1166 22.1 1089 1 US-08-180-195-36 Sequence 36, Appli
31 1166 22.1 1089 1 US-08-168-917-4 Sequence 4, Appli
32 1166 22.1 1089 1 US-08-477-329-36 Sequence 36, Appli
33 1166 22.1 1089 2 US-08-475-458-36 Sequence 4, Appli
34 1166 22.1 1089 2 US-08-460-510-4 Sequence 4, Appli
35 1166 22.1 1089 3 US-08-460-490-4 Sequence 36, Appli
36 1166 22.1 1089 3 US-08-980-400-36 Sequence 36, Appli
37 1166 22.1 1089 3 US-08-462-728-2 Sequence 2, Appli
38 1166 22.1 1089 3 US-09-583-459A-36 Sequence 36, Appli
39 1166 22.1 1089 3 US-09-583-210-36 Sequence 36, Appli
40 1166 22.1 1089 4 US-09-583-449A-36 Sequence 36, Appli
41 1166 22.1 1089 4 US-09-435-059-36 Sequence 36, Appli
42 1166 22.1 1089 4 US-08-461-917-2 Sequence 2, Appli
43 1166 22.1 1089 4 US-08-464-436-2 Sequence 2, Appli
44 1166 22.1 1089 5 PCT-US92-00730-4 Sequence 4, Appli
45 1166 22.1 1089 5 PCT-US92-00862-4 Sequence 4, Appli

ALIGNMENTS

RESULT 1
US-07-977-451-4
; Sequence 4, Application US/07977451
; Patent No. 5270458
; GENERAL INFORMATION:
; APPLICANT: Lemischka, Ihor R.
; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL
; TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ImClone Systems Incorporated
; STREET: 180 Varick Street
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10014
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/977,451
; FILING DATE: 19921119
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US UNASSIGNED
; FILING DATE: 12-NOV-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/906,397
; FILING DATE: 26-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US PCT/US92/05401
; FILING DATE: 26-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: TW 81102961
; FILING DATE: 15-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US PCT/US92/02750
; FILING DATE: 02-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/813,593
; FILING DATE: 24-DEC-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/793,065
; FILING DATE: 15-NOV-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/728,913
; FILING DATE: 28-JUN-1991
; PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/679,666
 FILING DATE: 02-APR-1991
 ATTORNEY/AGENT INFORMATION:
 NAME: Feit, Irving N.
 REGISTRATION NUMBER: 28,601
 REFERENCE/DOCKET NUMBER: LEM-3-7P
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 212-645-1405
 TELEFAX: 212-645-2054
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 993 amino acids
 TYPE: AMINO ACID
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-07-977-451-4

Query Match 100.0%; Score 5274; DB 1; Length 993;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 993; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPALARDAGTVPVLLVFSAMIFGTITNQDLPIKVLINHNKNDSSVGKSSSYPMVSESP 60
 Db 1 MPALARDAGTVPVLLVFSAMIFGTITNQDLPIKVLINHNKNDSSVGKSSSYPMVSESP 60
 QY 61 EDLGCALRQSSGTVEAAVVDVSAITLQVLVDAPGNISCLWPKHSLNCQPHFDL 120
 Db 61 EDLGCALRQSSGTVEAAVVDVSAITLQVLVDAPGNISCLWPKHSLNCQPHFDL 120
 QY 121 QNRGVSMVLKMTETQAGYLLFIQSEATNTILFTVSRNTLLYTLRRPYFRKMNQD 180
 Db 121 QNRGVSMVLKMTETQAGYLLFIQSEATNTILFTVSRNTLLYTLRRPYFRKMNQD 180
 QY 181 ALVCISESPEPIVEWVLCDSQGESCKEESPAVVKKEKVLHELFGTDIRCCARNELGRE 240
 Db 181 ALVCISESPEPIVEWVLCDSQGESCKEESPAVVKKEKVLHELFGTDIRCCARNELGRE 240
 QY 241 CTLRFTIDLNQTTLPQLFLKVGPELWIRCAVHVNHGFGLTWELNKALEEGNFEM 300
 Db 241 CTLRFTIDLNQTTLPQLFLKVGPELWIRCAVHVNHGFGLTWELNKALEEGNFEM 300
 QY 301 STYSTNRMTIRILFAVSSVARNDTGYTCCSSKHPQSQALVTIVGKGFNATNSSDYE 360
 Db 301 STYSTNRMTIRILFAVSSVARNDTGYTCCSSKHPQSQALVTIVGKGFNATNSSDYE 360
 QY 361 IDQYEEFCFSVRKAYPOIRCTWTFRSKPPCEQKGLDNGYSISKFNHKKHQCEYIFHA 420
 Db 361 IDQYEEFCFSVRKAYPOIRCTWTFRSKPPCEQKGLDNGYSISKFNHKKHQCEYIFHA 420
 QY 421 ENDDAQTKMFTNIRKPOVLAEASASQSCFSDGYPLPSWTWKCKSDKSPNCTEITE 480
 Db 421 ENDDAQTKMFTNIRKPOVLAEASASQSCFSDGYPLPSWTWKCKSDKSPNCTEITE 480
 QY 481 GVNRRKANRVQGWSSSTLNSEATLKGFLKCCAYNSLGTSCETILLNSPGPFPIQD 540
 Db 481 GVNRRKANRVQGWSSSTLNSEATLKGFLKCCAYNSLGTSCETILLNSPGPFPIQD 540
 QY 541 NISFYATIGVCLLPIVVLTLILCHYKKQRYESQOLQVQVTSDDNEFYFVDPREYED 600
 Db 541 NISFYATIGVCLLPIVVLTLILCHYKKQRYESQOLQVQVTSDDNEFYFVDPREYED 600
 QY 601 LKWEFFPRENLEFGKVLGSGAFGKVMNATAYGISTKTVSIOVAVKMLKEKADSSREALMS 660
 Db 601 LKWEFFPRENLEFGKVLGSGAFGKVMNATAYGISTKTVSIOVAVKMLKEKADSSREALMS 660
 QY 661 ELKMTOLGSHENIVNLGACTLSGPIYLIFEYCCYGLNLYLRSKREKTHRTWTEIFKE 720
 Db 661 ELKMTOLGSHENIVNLGACTLSGPIYLIFEYCCYGLNLYLRSKREKTHRTWTEIFKE 720
 QY 721 HNFSEFYTFOSHPSNMPGSGREVQIHDPDSQOISGLHGNFSHSEDEIEYENOKRLEEBEDL 780
 Db 721 HNFSEFYTFOSHPSNMPGSGREVQIHDPDSQOISGLHGNFSHSEDEIEYENOKRLEEBEDL 780

QY 781 NVLTPEDLICFAYQVAKGMEFLFKSCVHRDLAARNVLTGHKVVKICDFGLARDIMSDS 840
 Db 781 NVLTPEDLICFAYQVAKGMEFLFKSCVHRDLAARNVLTGHKVVKICDFGLARDIMSDS 840
 QY 841 NYVVRGNARLPVKWMAPELSEFEGITYIKSDVMSYGILLWEIFSLGVNPPGIPVDANFYK 900
 Db 841 NYVVRGNARLPVKWMAPELSEFEGITYIKSDVMSYGILLWEIFSLGVNPPGIPVDANFYK 900
 QY 901 LIQNGFKMDOPFYATEEIIYIMQSWAFDSRKRKPSFNLTSLFGCOLADAEAEAMYQNVDG 960
 Db 901 LIQNGFKMDOPFYATEEIIYIMQSWAFDSRKRKPSFNLTSLFGCOLADAEAEAMYQNVDG 960
 QY 961 RVSECPHTYQNRPPSRKEMDLGLLSPQAQVEDS 993
 Db 961 RVSECPHTYQNRPPSRKEMDLGLLSPQAQVEDS 993

RESULT 2
 US-08-252-517-4
 ; Sequence 4, Application US/08252517
 ; Patent No. 5548065
 ; GENERAL INFORMATION:
 ; APPLICANT: Lemischka, Ihor R.
 ; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL
 ; TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS
 ; NUMBER OF SEQUENCES: 10
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: ImClone Systems Incorporated
 ; STREET: 180 Varick Street
 ; CITY: New York
 ; STATE: New York
 ; COUNTRY: U.S.A.
 ; ZIP: 10014
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/252,517
 ; FILING DATE: 31-OCT-1994
 ; CLASSIFICATION: 530
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/977,451
 ; FILING DATE: 19-NOV-1992
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/906,397
 ; FILING DATE: 26-JUN-1992
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US PCT/US92/05401
 ; FILING DATE: 26-JUN-1992
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: TW 81102961
 ; FILING DATE: 15-APR-1992
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US PCT/US92/02750
 ; FILING DATE: 02-APR-1992
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/813,593
 ; FILING DATE: 24-DEC-1991
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/793,065
 ; FILING DATE: 15-NOV-1991
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/728,913
 ; FILING DATE: 28-JUN-1991
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/679,666
 ; FILING DATE: 02-APR-1991
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Feit, Irving N.
 ; REGISTRATION NUMBER: 28,601
 ; REFERENCE/DOCKET NUMBER: LEM-3-7P

TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-645-1405
TELEFAX: 212-645-2054
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 993 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-252-517-4

Query Match 100.0%; Score 5274; DB 1; Length 993;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 993; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MPALARDAGTVPLLVVFSAMIFGTTNQDLPVKVCLINHKNDSSVGKSSSYPMVSESP 60
Db 1 MPALARDAGTVPLLVVFSAMIFGTTNQDLPVKVCLINHKNDSSVGKSSSYPMVSESP 60

Qy 61 EDLGCALRPOSSGTYEAAVEVDVSASITLQVLVDAPGNISCLWFKHSSINCOHPFDL 120
Db 61 EDLGCALRPOSSGTYEAAVEVDVSASITLQVLVDAPGNISCLWFKHSSINCOHPFDL 120

Qy 121 QNRGVSMVILKMTQAGEYLLFTIOSEATNYTILFTVSIIRNTLYTLRRPYFRKMNQD 180
Db 121 QNRGVSMVILKMTQAGEYLLFTIOSEATNYTILFTVSIIRNTLYTLRRPYFRKMNQD 180

Qy 181 ALVCISESVPEPIVEWVLCDQSGESCKEESPAVVKKEKVLHFGTDIRCCARNELGRE 240
Db 181 ALVCISESVPEPIVEWVLCDQSGESCKEESPAVVKKEKVLHFGTDIRCCARNELGRE 240

Qy 241 CTRLTFTDLNQTPTTLPOLFKVGEPLWIRKAVVNHGGLTWELNKALEEGNYFEM 300
Db 241 CTRLTFTDLNQTPTTLPOLFKVGEPLWIRKAVVNHGGLTWELNKALEEGNYFEM 300

Qy 301 STYSTNRTMIRILFAFVSSVARNDGYTCSSRHPSOSALVTIVKGFINATNSSDEYE 360
Db 301 STYSTNRTMIRILFAFVSSVARNDGYTCSSRHPSOSALVTIVKGFINATNSSDEYE 360

Qy 361 IDQYEEFCFSVRKAYPOIRCTWFSRKSFFCEQKGLDNGYSISKFNHKGHPGEYIFHA 420
Db 361 IDQYEEFCFSVRKAYPOIRCTWFSRKSFFCEQKGLDNGYSISKFNHKGHPGEYIFHA 420

Qy 421 ENDDAQFTKMTLNIIRKQVLAESAQSCFSDGYPLPWNWTKCKSDKSPNCTEITE 480
Db 421 ENDDAQFTKMTLNIIRKQVLAESAQSCFSDGYPLPWNWTKCKSDKSPNCTEITE 480

Qy 481 GVMNRKANRKYFGQWVSSSTLNMSEAIGFLVKCCAYNSLGTSCETILLNSPGPPPTQD 540
Db 481 GVMNRKANRKYFGQWVSSSTLNMSEAIGFLVKCCAYNSLGTSCETILLNSPGPPPTQD 540

Qy 541 NISFVATIGVCLLFTVLTLLIHKYKQFRYESQLOMVQVGTSSDNEYFYVDREYEYD 600
Db 541 NISFVATIGVCLLFTVLTLLIHKYKQFRYESQLOMVQVGTSSDNEYFYVDREYEYD 600

Qy 601 LKWEFPRENLEFGKVLGSGAFGKVNATAYGISKRTGSIQVAVKMLKEKADSSEREALMS 660
Db 601 LKWEFPRENLEFGKVLGSGAFGKVNATAYGISKRTGSIQVAVKMLKEKADSSEREALMS 660

Qy 661 ELKMTQLGSHENIVNLGACTLSPIYLIPEYCCYGDLLNLYLRKREKFRHTWTETPKE 720
Db 661 ELKMTQLGSHENIVNLGACTLSPIYLIPEYCCYGDLLNLYLRKREKFRHTWTETPKE 720

Qy 721 HNFSPYPTFQSHPNSSMPGSRVQIHPDSQDQISGLHGNFSFSEDEIYEENOKRLEEDL 780
Db 721 HNFSPYPTFQSHPNSSMPGSRVQIHPDSQDQISGLHGNFSFSEDEIYEENOKRLEEDL 780

Qy 781 NVLTFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVVKICDFGLARDIMSDS 840
Db 781 NVLTFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVVKICDFGLARDIMSDS 840

Qy 841 NYVVRGNARLPVKWAPESLPEGITYTIKSDVMSYGILLWEIFSLGVNFPYGPVDPANFYK 900

Db 841 NYVVRGNARLPVKWAPESLPEGITYTIKSDVMSYGILLWEIFSLGVNFPYGPVDPANFYK 900

Qy 901 LIQNGFMDQFPYATEEIIIMQSCWAFDSKRSPFNLTSLFLGCOLADAEAMTYQNDG 960

Db 901 LIQNGFMDQFPYATEEIIIMQSCWAFDSKRSPFNLTSLFLGCOLADAEAMTYQNDG 960

Qy 961 RVSECPHTYQNRPPFSREMDLGLLSPOAQVEDS 993

Db 961 RVSECPHTYQNRPPFSREMDLGLLSPOAQVEDS 993

RESULT 3
US-07-906-397A-4
; Sequence 4, Application US/07906397A
; Patent No. 5621090
; GENERAL INFORMATION:
; APPLICANT: Lemischka, Ihor R.
; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL
; TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: IMCLONE SYSTEMS INCORPORATED
; STREET: 180 VARICK STREET
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: U.S.A.
; ZIP: 10014
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/906.397A
; FILING DATE: 19920626
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/813,593
; FILING DATE: 24-DEC-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/793,065
; FILING DATE: 15-NOV-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/728,913
; FILING DATE: 28-JUN-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/679,666
; FILING DATE: 02-APR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Feit, Irving N.
; REGISTRATION NUMBER: 28,601
; REFERENCE/DOCKET NUMBER: LEM-3-PPPPPP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-645-1405
; TELEFAX: 212-645-2054
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 993 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-906-397A-4

Query Match 100.0%; Score 5274; DB 1; Length 993;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 993; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MPALARDAGTVPLLVVFSAMIFGTTNQDLPVKVCLINHKNDSSVGKSSSYPMVSESP 60

Db 1 MPALARDAGTVPLLVVFSAMIFGTTNQDLPVKVCLINHKNDSSVGKSSSYPMVSESP 60

Qy 61 EDLGCALRPOSSGTYEAAVEVDVSASITLQVLVDAPGNISCLWFKHSSINCOHPFDL 120

61 EDLGCALPQSSGTVYAAAEVDVVSASITLQVLVDAPGNISCLWVFKHSSLNCPHFDL 120
121 QNRGVSVMLKMTQAGEYLLFIQSEATNTYILFTVSIIRNTLLYTLRRPYFRMENQD 180
121 QNRGVSVMLKMTQAGEYLLFIQSEATNTYILFTVSIIRNTLLYTLRRPYFRMENQD 180
181 ALVCISEVPPIVWVLCDSGSCKEESPAVVKKEKVLHELFGTDIRCCARNELGRE 240
181 ALVCISEVPPIVWVLCDSGSCKEESPAVVKKEKVLHELFGTDIRCCARNELGRE 240
241 CTRFTIDLNTQPTTLTLPFLKVGCEPLWIRCKAVHVNHGFLTWELNKALEEGNYFEM 300
241 CTRFTIDLNTQPTTLTLPFLKVGCEPLWIRCKAVHVNHGFLTWELNKALEEGNYFEM 300
301 STYSTNRTMIRILFAFVSSVARNDGYTTCSSSKHPSQSALVTIIVGKGFINATNSSEDEY 360
301 STYSTNRTMIRILFAFVSSVARNDGYTTCSSSKHPSQSALVTIIVGKGFINATNSSEDEY 360
361 IDQYEEPCFSVRFKAYPOIRCTWTFSRKSFPCEQKGLONGYISIKFCNHHKQPGYIIFHA 420
361 IDQYEEPCFSVRFKAYPOIRCTWTFSRKSFPCEQKGLONGYISIKFCNHHKQPGYIIFHA 420
421 ENDDAQFTKMTLNTIRRRPQVLAESAQSCFSDGYPLPSWTWKKCSDKSPNCTEETITE 480
421 ENDDAQFTKMTLNTIRRRPQVLAESAQSCFSDGYPLPSWTWKKCSDKSPNCTEETITE 480
481 GVNWRKANRKFVGQWVSSSTLNNBAIKGFLVKCAVNSLGTSCETILLNSPGPPFFIQD 540
481 GVNWRKANRKFVGQWVSSSTLNNBAIKGFLVKCAVNSLGTSCETILLNSPGPPFFIQD 540
541 NISFYATIGVCLLFIVLTLTLLCHYKQFRYESQLQWQVGTSSDNEYFYVDREYED 600
541 NISFYATIGVCLLFIVLTLTLLCHYKQFRYESQLQWQVGTSSDNEYFYVDREYED 600
601 LKWEPPRENLEFGKVLGSGAFQKVNATAYGSKTGVSIQVAVNMLKEKADSSEREALMS 660
601 LKWEPPRENLEFGKVLGSGAFQKVNATAYGSKTGVSIQVAVNMLKEKADSSEREALMS 660
661 ELKMTQLGSHENIVNLGACTLSGPYILFIYCCVGDLLNLRKREKFRHTWTEIPKE 720
661 ELKMTQLGSHENIVNLGACTLSGPYILFIYCCVGDLLNLRKREKFRHTWTEIPKE 720
721 HNFSPYPTFQSHNPSMPSREVRQIHPSDQISGLHNSFHSEDEIEYENQKRLSEEDL 780
721 HNFSPYPTFQSHNPSMPSREVRQIHPSDQISGLHNSFHSEDEIEYENQKRLSEEDL 780
781 NVLTFEDLLCFAYQAVAKGMEFLFKSCVHRDLAARNVLTGKVKVVICDFGLARDIMSDS 840
781 NVLTFEDLLCFAYQAVAKGMEFLFKSCVHRDLAARNVLTGKVKVVICDFGLARDIMSDS 840
841 NVVVRGNARLPVKWAPESLFGIYTIKSDVNSYGILLWEIFSLGWNYPYGPVDFANFYK 900
841 NVVVRGNARLPVKWAPESLFGIYTIKSDVNSYGILLWEIFSLGWNYPYGPVDFANFYK 900
901 LIQNGFKMDQPFYATEEIIYIMQSWAFDSRKRPFNPNTSLFGCOLADAEAMYNQVNDG 960
901 LIQNGFKMDQPFYATEEIIYIMQSWAFDSRKRPFNPNTSLFGCOLADAEAMYNQVNDG 960
961 RVSECPHYQNRPRPSREMDLGLSPQAQVEDS 993
961 RVSECPHYQNRPRPSREMDLGLSPQAQVEDS 993

RESULT 4

US-08-601-891-4
; Sequence 4, Application US/08601891
; Patent No. 5747651
; GENERAL INFORMATION:

; APPLICANT: Lemischka, Ihor R.
; TITLE OF INVENTION: TOPIOTENT HEMATOPOIETIC STEM CELL
; TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:

ADDRESSEE: ImClone Systems Incorporated
STREET: 180 Varick Street
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10014
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/601,891
FILING DATE: 15-FEB-1996
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/977,451
FILING DATE: 19-NOV-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/906,397
FILING DATE: 26-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US PCT/US92/05401
FILING DATE: 26-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: TW 81102961
FILING DATE: 15-APR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US PCT/US92/02750
FILING DATE: 02-APR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/813,593
FILING DATE: 24-DEC-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/793,065
FILING DATE: 15-NOV-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/728,913
FILING DATE: 28-JUN-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/679,666
FILING DATE: 02-APR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Feit, Irving N.
REGISTRATION NUMBER: 28,601
REFERENCE/DOCKET NUMBER: LEM-3-7P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-645-1405
TELEFAX: 212-645-2054
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 993 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-601-891-4

Query Match 100.0%; Score 5274; DB 1; Length 993;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 993; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MPALARDAGTVPLLVVFSAMIFGTITNQDLPVVKVILNHNKNDSSVGVKSSSYPMVSESP 60
DB 1 MPALARDAGTVPLLVVFSAMIFGTITNQDLPVVKVILNHNKNDSSVGVKSSSYPMVSESP 60
QY 61 EDLGCALPQSSGTVYAAAEVDVVSASITLQVLVDAPGNISCLWVFKHSSLNCPHFDL 120
DB 61 EDLGCALPQSSGTVYAAAEVDVVSASITLQVLVDAPGNISCLWVFKHSSLNCPHFDL 120
QY 121 QNRGVSVMLKMTQAGEYLLFIQSEATNTYILFTVSIIRNTLLYTLRRPYFRMENQD 180
DB 121 QNRGVSVMLKMTQAGEYLLFIQSEATNTYILFTVSIIRNTLLYTLRRPYFRMENQD 180

181 ALVCISESVPEPIVWVLCDSQSGCKEESPAVVKKEKVLHFGTDIRCCARNELGRE 240
 181 ALVCISESVPEPIVWVLCDSQSGCKEESPAVVKKEKVLHFGTDIRCCARNELGRE 240
 241 CTRLTIDLNQTPQTTLPQLFKVGEPLWIRKAVHNVHGGFGLTWELNKALEBGNFYFEM 300
 241 CTRLTIDLNQTPQTTLPQLFKVGEPLWIRKAVHNVHGGFGLTWELNKALEBGNFYFEM 300
 301 STYSTNRTMIRILFAFVSSVARNDTGYTCSSSKHPSQSALVTIVGGINATNSSDEYE 360
 301 STYSTNRTMIRILFAFVSSVARNDTGYTCSSSKHPSQSALVTIVGGINATNSSDEYE 360
 361 IDOVEECFVSVPKAYQIRCTWFSKSPCEQKGLDNGYSISKFCNKHKHOPEYIFPHA 420
 361 IDOVEECFVSVPKAYQIRCTWFSKSPCEQKGLDNGYSISKFCNKHKHOPEYIFPHA 420
 421 ENDDAQTKFTLNIRKPKQVLAASASQSCFSDGYPLPSWTWKCKSDKSPNCTEITE 480
 421 ENDDAQTKFTLNIRKPKQVLAASASQSCFSDGYPLPSWTWKCKSDKSPNCTEITE 480
 481 GVNRRKANRVFGQWVSSSTLNSEAIKGLVKCCAYNSLGTSCETILLNSPGFPFIQD 540
 481 GVNRRKANRVFGQWVSSSTLNSEAIKGLVKCCAYNSLGTSCETILLNSPGFPFIQD 540
 541 NISFYATIGVCLLFIIVVLTLLIHKYKQRYESQLOMVQVGTSSDNEYFYVDREYEYD 600
 541 NISFYATIGVCLLFIIVVLTLLIHKYKQRYESQLOMVQVGTSSDNEYFYVDREYEYD 600
 601 LKWEFFPRENLFGKVLGSGAFKVMNATAYGISTGVSIOAVKMLKADSSEREALMS 660
 601 LKWEFFPRENLFGKVLGSGAFKVMNATAYGISTGVSIOAVKMLKADSSEREALMS 660
 661 ELQWMTQLGSHENIVNLGACTLSGPTLYLFEYCCYGDLLNLYLSKREKHFRTWTEIFKE 720
 661 ELQWMTQLGSHENIVNLGACTLSGPTLYLFEYCCYGDLLNLYLSKREKHFRTWTEIFKE 720
 721 HNFSEFPTFOSHNSMPGREGVQIHDPDSQISGLHNSHSEDEIEYENQKLEEEEDL 780
 721 HNFSEFPTFOSHNSMPGREGVQIHDPDSQISGLHNSHSEDEIEYENQKLEEEEDL 780
 781 NVLTFFDLCLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVKVVICDFGLARDIMSDS 840
 781 NVLTFFDLCLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVKVVICDFGLARDIMSDS 840
 841 NYVVRGNARLPVKWMAPESLFEGITYTKSDVWSYGILLWEIFSLGVNPNYPGIPVDANFYK 900
 841 NYVVRGNARLPVKWMAPESLFEGITYTKSDVWSYGILLWEIFSLGVNPNYPGIPVDANFYK 900
 901 LIONGFMDQPPFYATEIYIIMOSCAFDKRPSPFNLTSLGCOLADAEAMYNQVNDG 960
 901 LIONGFMDQPPFYATEIYIIMOSCAFDKRPSPFNLTSLGCOLADAEAMYNQVNDG 960
 961 RVSECPHTYQNRPPFSEMDLGLLSPQAQVEDS 993
 961 RVSECPHTYQNRPPFSEMDLGLLSPQAQVEDS 993

RESULT 5
 US-09-021-324-4
 ; Sequence 4, Application US/09021324
 ; Patent No. 5912133
 ; GENERAL INFORMATION:
 ; APPLICANT: Lemischka, Ihor R.
 ; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL
 ; TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS
 ; NUMBER OF SEQUENCES: 10
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: ImClone Systems Incorporated
 ; STREET: 180 Varick Street
 ; CITY: New York
 ; STATE: New York
 ; COUNTRY: U.S.A.
 ; ZIP: 10014

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/021,324
 FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/07/977,451
 FILING DATE: 1992-11-19
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/906,397
 FILING DATE: 26-JUN-1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US PCT/US92/05401
 FILING DATE: 26-JUN-1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: TW 81102961
 FILING DATE: 15-APR-1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US PCT/US92/02750
 FILING DATE: 02-APR-1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/813,593
 FILING DATE: 24-DEC-1991
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/793,065
 FILING DATE: 15-NOV-1991
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/728,913
 FILING DATE: 28-JUN-1991
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/679,666
 FILING DATE: 02-APR-1991
 ATTORNEY/AGENT INFORMATION:
 NAME: Feit, Irving N.
 REGISTRATION NUMBER: 28,601
 REFERENCE/DOCKET NUMBER: LEM-3-7P
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 212-645-1405
 TELEFAX: 212-645-2054
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 993 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-09-021-324-4

Query Match 100.0%; Score 5274; DB 2; Length 993;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 993; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MPALARDAGTVPVLLVVFSAIFGTITNQDLFVKVCLINHKNDSSVKGSSSYPMVSESP 60
 DB 1 MPALARDAGTVPVLLVVFSAIFGTITNQDLFVKVCLINHKNDSSVKGSSSYPMVSESP 60
 QY 61 EDLGCALRPOSSGTVYEAATAVEVDVSASITLQVLVDAPGNISCLWPKHSSLNCQPHFDL 120
 DB 61 EDLGCALRPOSSGTVYEAATAVEVDVSASITLQVLVDAPGNISCLWPKHSSLNCQPHFDL 120
 QY 121 QNRGVVSWILKMTETQAGEYLLFIOSEATNYTLFTVSRNTLLYTLRPPYFRKMNQD 180
 DB 121 QNRGVVSWILKMTETQAGEYLLFIOSEATNYTLFTVSRNTLLYTLRPPYFRKMNQD 180
 QY 181 ALVCISESVPEPIVWVLCDSQSGCKEESPAVVKKEKVLHFGTDIRCCARNELGRE 240
 DB 181 ALVCISESVPEPIVWVLCDSQSGCKEESPAVVKKEKVLHFGTDIRCCARNELGRE 240
 QY 241 CTRLTIDLNQTPQTTLPQLFKVGEPLWIRKAVHNVHGGFGLTWELNKALEBGNFYFEM 300

Db 241 CTRLTIDLNTQPTTLTLPQLFLKVGELPLIRCAVHVNHGFLGTWELNKALEEGNYFEM 300
Qy 301 STYSTNRTRMIRILFAFVSSVARNDTGYYTCSSSKHPSQSALVTIVGKGFNATNSSEDEY 360
Db 301 STYSTNRTRMIRILFAFVSSVARNDTGYYTCSSSKHPSQSALVTIVGKGFNATNSSEDEY 360
Qy 361 IDQYEFECFSVRPKAYPQIRCTWTFRSKSPCCQKGLDNGYSISKFCNKHQPGGEYIFHA 420
Db 361 IDQYEFECFSVRPKAYPQIRCTWTFRSKSPCCQKGLDNGYSISKFCNKHQPGGEYIFHA 420
Qy 421 ENDDAQFTKMTLNIIRKPKQVLAESAASQASCFSDGYPLPSWTWKCKSDKSPNCTEITE 480
Db 421 ENDDAQFTKMTLNIIRKPKQVLAESAASQASCFSDGYPLPSWTWKCKSDKSPNCTEITE 480
Qy 481 GWNRRKANRKFQGVSSSTLNNSSEAIKGLVKKCCAYNSLGTSCETILLNSPGPPFIQD 540
Db 481 GWNRRKANRKFQGVSSSTLNNSSEAIKGLVKKCCAYNSLGTSCETILLNSPGPPFIQD 540
Qy 541 NISFYATIGVCLLFIVVLTLLI CHYKQKQPRYSQLOMVOVTGSSDNEYFYVDPREYED 600
Db 541 NISFYATIGVCLLFIVVLTLLI CHYKQKQPRYSQLOMVOVTGSSDNEYFYVDPREYED 600
Qy 601 LWKEFPRENLEFGKVLGSGAFGKVMNATAYGISKTVGSIQVAVKMLKEKADSSERREALMS 660
Db 601 LWKEFPRENLEFGKVLGSGAFGKVMNATAYGISKTVGSIQVAVKMLKEKADSSERREALMS 660
Qy 661 ELKQMTQOLGSHENIVNLLGACTLSGPIYLI FEYCCYGDLLNLYLSRKREKPHRTWTETPKE 720
Db 661 ELKQMTQOLGSHENIVNLLGACTLSGPIYLI FEYCCYGDLLNLYLSRKREKPHRTWTETPKE 720
Qy 721 HNFSPYPTFQSHNSMPSGSRVQIHPDSQIISGLHGNFSHSEDEIYEYENOKRLEEBEDL 780
Db 721 HNFSPYPTFQSHNSMPSGSRVQIHPDSQIISGLHGNFSHSEDEIYEYENOKRLEEBEDL 780
Qy 781 NVLTFFEDLLCFAYQVAKGMFLEFKSCVHRDLAARNVLVTHGKVKICDFGLARDIMSDS 840
Db 781 NVLTFFEDLLCFAYQVAKGMFLEFKSCVHRDLAARNVLVTHGKVKICDFGLARDIMSDS 840
Qy 841 NYVVRGNARLPVKWMAPESEFEGITTIKSDVWSYGLLLWEIFSLGVNPNYPGIPVDANFYK 900
Db 841 NYVVRGNARLPVKWMAPESEFEGITTIKSDVWSYGLLLWEIFSLGVNPNYPGIPVDANFYK 900
Qy 901 LIQNGFKMDOPFYATEIYIIMOSCAFDKSRKPSFNLTSFLGCQADAEAEAMYQNVGD 960
Db 901 LIQNGFKMDOPFYATEIYIIMOSCAFDKSRKPSFNLTSFLGCQADAEAEAMYQNVGD 960
Qy 961 RVSECPHTYQNRPPFSREMDLGLLSPQAQVEDS 993
Db 961 RVSECPHTYQNRPPFSREMDLGLLSPQAQVEDS 993

RESULT 6

PCT-US92-09893-4
; Sequence 4, Application PC/TUS9209893
; GENERAL INFORMATION:
; APPLICANT: Lemischka, Ihor R.
; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL
; TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ImClone Systems Incorporated
; STREET: 180 Varick Street
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10014
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/09893

FILING DATE: 19921116
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Felt, Irving N.
REGISTRATION NUMBER: 28,601
REFERENCE/DOCKET NUMBER: LEM-3-7PT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-645-1405
TELEFAX: 212-645-2054
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 993 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US92-09893-4
Query Match 100.0%; Score 5274; DB 5; Length 993;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 993; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MPALARDAGTVPLLVVFSAIFGTITNQDLPVTKCVLINHKNNDSSVGKSSSYPMVSESP 60
Db 1 MPALARDAGTVPLLVVFSAIFGTITNQDLPVTKCVLINHKNNDSSVGKSSSYPMVSESP 60
Qy 61 EDLGCALRPOSSGTVYEAFAVEVDVSASITLQVLVDAPGNI SCLWVPKHSLSNCPHFDL 120
Db 61 EDLGCALRPOSSGTVYEAFAVEVDVSASITLQVLVDAPGNI SCLWVPKHSLSNCPHFDL 120
Qy 121 QNRGVVSWTLKMTQAGEYLLFIQSEATNYTILFTVSRNTLLYTLRRPYPRKMNQD 180
Db 121 QNRGVVSWTLKMTQAGEYLLFIQSEATNYTILFTVSRNTLLYTLRRPYPRKMNQD 180
Qy 181 ALVCISESPEPIVWVLCDSQGESCKEESPAVVKKEKVLHFLGTDIRCCARNELGRE 240
Db 181 ALVCISESPEPIVWVLCDSQGESCKEESPAVVKKEKVLHFLGTDIRCCARNELGRE 240
Qy 241 CTRLTIDLNTQPTTLTLPQLFLKVGELPLIRCAVHVNHGFLGTWELNKALEEGNYFEM 300
Db 241 CTRLTIDLNTQPTTLTLPQLFLKVGELPLIRCAVHVNHGFLGTWELNKALEEGNYFEM 300
Qy 301 STYSTNRTRMIRILFAFVSSVARNDTGYYTCSSSKHPSQSALVTIVGKGFNATNSSEDEY 360
Db 301 STYSTNRTRMIRILFAFVSSVARNDTGYYTCSSSKHPSQSALVTIVGKGFNATNSSEDEY 360
Qy 361 IDQYEFECFSVRPKAYPQIRCTWTFRSKSPCCQKGLDNGYSISKFCNKHQPGGEYIFHA 420
Db 361 IDQYEFECFSVRPKAYPQIRCTWTFRSKSPCCQKGLDNGYSISKFCNKHQPGGEYIFHA 420
Qy 421 ENDDAQFTKMTLNIIRKPKQVLAESAASQASCFSDGYPLPSWTWKCKSDKSPNCTEITE 480
Db 421 ENDDAQFTKMTLNIIRKPKQVLAESAASQASCFSDGYPLPSWTWKCKSDKSPNCTEITE 480
Qy 481 GWNRRKANRKFQGVSSSTLNNSSEAIKGLVKKCCAYNSLGTSCETILLNSPGPPFIQD 540
Db 481 GWNRRKANRKFQGVSSSTLNNSSEAIKGLVKKCCAYNSLGTSCETILLNSPGPPFIQD 540
Qy 541 NISFYATIGVCLLFIVVLTLLI CHYKQKQPRYSQLOMVOVTGSSDNEYFYVDPREYED 600
Db 541 NISFYATIGVCLLFIVVLTLLI CHYKQKQPRYSQLOMVOVTGSSDNEYFYVDPREYED 600
Qy 601 LWKEFPRENLEFGKVLGSGAFGKVMNATAYGISKTVGSIQVAVKMLKEKADSSERREALMS 660
Db 601 LWKEFPRENLEFGKVLGSGAFGKVMNATAYGISKTVGSIQVAVKMLKEKADSSERREALMS 660
Qy 661 ELKQMTQOLGSHENIVNLLGACTLSGPIYLI FEYCCYGDLLNLYLSRKREKPHRTWTETPKE 720
Db 661 ELKQMTQOLGSHENIVNLLGACTLSGPIYLI FEYCCYGDLLNLYLSRKREKPHRTWTETPKE 720
Qy 721 HNFSPYPTFQSHNSMPSGSRVQIHPDSQIISGLHGNFSHSEDEIYEYENOKRLEEBEDL 780
Db 721 HNFSPYPTFQSHNSMPSGSRVQIHPDSQIISGLHGNFSHSEDEIYEYENOKRLEEBEDL 780

QY 781 NVLTFFDLCCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVKVVICDFGLARDIMSDS 840
 DB 781 NVLTFFDLCCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVKVVICDFGLARDIMSDS 840
 QY 841 NYVVRGNARLPVKWMAPELSEFEGYTIKSDVWSYGILLWEIFSLGVNYPGIPVDANFYK 900
 DB 841 NYVVRGNARLPVKWMAPELSEFEGYTIKSDVWSYGILLWEIFSLGVNYPGIPVDANFYK 900
 QY 901 LIQNGFKMDQPFYATEEIIYIMQSCWAFDSKRPSPNLTSLFCOLADAEAMYNQVNDG 960
 DB 901 LIQNGFKMDQPFYATEEIIYIMQSCWAFDSKRPSPNLTSLFCOLADAEAMYNQVNDG 960
 QY 961 RVSECPHTYQNRPRFSPREMDLGLLSPQAQVEDS 993
 DB 961 RVSECPHTYQNRPRFSPREMDLGLLSPQAQVEDS 993

RESULT 7
 PCT-US92-05401-4
 ; Sequence 4, Application PC/TUS9205401
 ; GENERAL INFORMATION:
 ; APPLICANT: Lemischka, Ihor R.
 ; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL
 ; TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS
 ; NUMBER OF SEQUENCES: 6
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: INCLONE SYSTEMS INCORPORATED
 ; STREET: 180 VARICK STREET
 ; CITY: NEW YORK
 ; STATE: NEW YORK
 ; COUNTRY: U.S.A.
 ; ZIP: 10014
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/US92/05401
 ; FILING DATE: 19920626
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Feit, Irving N.
 ; REGISTRATION NUMBER: 28,601
 ; REFERENCE/DOCKET NUMBER: LEM-3-PPPPPT
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 212-645-1405
 ; TELEFAX: 212-645-2054
 ; INFORMATION FOR SEQ ID NO: 4:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1160 amino acids
 ; TYPE: AMINO ACID
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; PCT-US92-05401-4

Query Match 100.0%; Score 5274; DB 5; Length 1160;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 993; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPALADAGTVPLLVFVSAMIFGTITNQDLPVKCVLINHKNDSSVGVKSSSYPMVWSESP 60
 DB 20 MPALADAGTVPLLVFVSAMIFGTITNQDLPVKCVLINHKNDSSVGVKSSSYPMVWSESP 79
 QY 61 EDLGCALRPOSSGTVEAAAVEVDVSASITLQVLVDAPGNISCLWVFKHSLNCQPHFDL 120
 DB 80 EDLGCALRPOSSGTVEAAAVEVDVSASITLQVLVDAPGNISCLWVFKHSLNCQPHFDL 139
 QY 121 QNRGVVSWILKMTQAGEYLLFIQSEATNYTLFTVSRINTLLYTLRPPYFRKMNQD 180
 DB 140 QNRGVVSWILKMTQAGEYLLFIQSEATNYTLFTVSRINTLLYTLRPPYFRKMNQD 199
 QY 181 ALVCISESPPEIIVENVLCDQGESCKEESPAVVKKEKVLHELFGTDIRCCARNELGRE 240

DB 200 ALVCISESPPEIIVENVLCDQGESCKEESPAVVKKEKVLHELFGTDIRCCARNELGRE 259
 QY 241 CTRLFTIDLNTPTTLPOLFLKVGEPLEWIRCKAVVNHGFLGTWELENKALEEGNYFEM 300
 DB 260 CTRLFTIDLNTPTTLPOLFLKVGEPLEWIRCKAVVNHGFLGTWELENKALEEGNYFEM 319
 QY 301 STYSTNRTMIRILFAFVSSVARNDTYTTCSSSKHPSQSALVTIIVKGFINATNSEDYE 360
 DB 320 STYSTNRTMIRILFAFVSSVARNDTYTTCSSSKHPSQSALVTIIVKGFINATNSEDYE 379
 QY 361 IDQYEEFCFSVRFKAYPOIRCTWTFSRKSFPCEQKGLDNGYSISKPCNHKHQPGVIFHA 420
 DB 380 IDQYEEFCFSVRFKAYPOIRCTWTFSRKSFPCEQKGLDNGYSISKPCNHKHQPGVIFHA 439
 QY 421 ENDDAQFTKMTLNIIRKPOVLAESASQASCFSDGYPLPSMTWKCKDSCNCTEEITE 480
 DB 440 ENDDAQFTKMTLNIIRKPOVLAESASQASCFSDGYPLPSMTWKCKDSCNCTEEITE 499
 QY 481 GVMRKANRKFVGQVSSSTLNMSEAINGFLVKCCAYNSLGTSCETILLNSGPPPIOD 540
 DB 500 GVMRKANRKFVGQVSSSTLNMSEAINGFLVKCCAYNSLGTSCETILLNSGPPPIOD 559
 QY 541 NISFYATIGVCLLFIVLTLIICHYKQFQYESOLOMVQVGTSSDNEFYVDFREYED 600
 DB 560 NISFYATIGVCLLFIVLTLIICHYKQFQYESOLOMVQVGTSSDNEFYVDFREYED 619
 QY 601 LKWEFPRENLEFGKVLGSGAFKVMNATAYIGISKTGVSIOQAVKMLKADSSEREALMS 660
 DB 620 LKWEFPRENLEFGKVLGSGAFKVMNATAYIGISKTGVSIOQAVKMLKADSSEREALMS 679
 QY 661 ELKMTQLGSHENIVNLLGACTLSGPIYLIPEYCCYGLLNLRSKRKFKHRTWTEIPKE 720
 DB 680 ELKMTQLGSHENIVNLLGACTLSGPIYLIPEYCCYGLLNLRSKRKFKHRTWTEIPKE 739
 QY 721 HNFSFYTFQSHNSMPGSRVQIHPSDDOISGLHGNFSHSEDEIYEYENOKRLEEDL 780
 DB 740 HNFSFYTFQSHNSMPGSRVQIHPSDDOISGLHGNFSHSEDEIYEYENOKRLEEDL 799
 QY 781 NVLTFFDLCCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVKVVICDFGLARDIMSDS 840
 DB 800 NVLTFFDLCCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVKVVICDFGLARDIMSDS 859
 QY 841 NYVVRGNARLPVKWMAPELSEFEGYTIKSDVWSYGILLWEIFSLGVNYPGIPVDANFYK 900
 DB 860 NYVVRGNARLPVKWMAPELSEFEGYTIKSDVWSYGILLWEIFSLGVNYPGIPVDANFYK 919
 QY 901 LIQNGFKMDQPFYATEEIIYIMQSCWAFDSKRPSPNLTSLFCOLADAEAMYNQVNDG 960
 DB 920 LIQNGFKMDQPFYATEEIIYIMQSCWAFDSKRPSPNLTSLFCOLADAEAMYNQVNDG 979
 QY 961 RVSECPHTYQNRPRFSPREMDLGLLSPQAQVEDS 993
 DB 980 RVSECPHTYQNRPRFSPREMDLGLLSPQAQVEDS 1012

RESULT 8
 US-08-222-299-4
 ; Sequence 4, Application US/08222299
 ; Patent No. 563538
 ; GENERAL INFORMATION:
 ; APPLICANT: Bennett, Brian D.
 ; APPLICANT: Broz, Susan D.
 ; APPLICANT: Matthews, William
 ; APPLICANT: Zeigler, Francis C.
 ; TITLE OF INVENTION: AGONIST ANTIBODIES AGAINST THE FLK2/FLT3 RECEPTOR AND USES TH
 ; NUMBER OF SEQUENCES: 4
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Genentech, Inc.
 ; STREET: 460 Point San Bruno Blvd
 ; CITY: South San Francisco
 ; STATE: California
 ; COUNTRY: USA


```

Query Match      99.4%; Score 5240; DB 2; Length 993;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 987; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 MPALARDAGTVPPLLVVFSAMIFGTITNQDLFVKVCLINHKNDSSVKGSSSYPMVSESP 60
DB 1 MPALARDGGQLPLLVVFSAMIFGTITNQDLFVKVCLINHKNDSSVKGSSSYPMVSESP 60
QY 61 EDLGCALRPOSSGTVYEAADVVSASITLQVLVDAPGNISCLWVFKHSSLNCQPHFDL 120
DB 61 EDLGCALRPOSSGTVYEAADVVSASITLQVLVDAPGNISCLWVFKHSSLNCQPHFDL 120
QY 121 QNRGVWSMWILKMTETQAGEYLLFIOSEATNTYTLFTVSRNTLLYTLRRPYFRMENQD 180
DB 121 QNRGVWSMWILKMTETQAGEYLLFIOSEATNTYTLFTVSRNTLLYTLRRPYFRMENQD 180
QY 181 ALVCISESVPEPIVWVLCDSQGESCKEESPAVVKKEKVLHFGTDIRCCARNELGRE 240
DB 181 ALVCISESVPEPIVWVLCDSQGESCKEESPAVVKKEKVLHFGTDIRCCARNELGRE 240
QY 241 CTRFLTIDLNOTPOTTLQPLFLKVGEPWIRCKAVHNVHGFGLTWELKALEEGNYFEM 300
DB 241 CTRFLTIDLNOTPOTTLQPLFLKVGEPWIRCKAVHNVHGFGLTWELKALEEGNYFEM 300
QY 301 STYSTNRMTIRILFAPVSSVARNDTGYTTCSSSKHPSQSALVTIVGKGFINATNSSDEYE 360
DB 301 STYSTNRMTIRILFAPVSSVARNDTGYTTCSSSKHPSQSALVTIVGKGFINATNSSDEYE 360
QY 361 IDQYEEFCFSVRKAYPOIRCTWTFRSKSPCEQKGLDNGYSISKFCNKHQPGYEYIFHA 420
DB 361 IDQYEEFCFSVRKAYPOIRCTWTFRSKSPCEQKGLDNGYSISKFCNKHQPGYEYIFHA 420
QY 421 ENDDAQTMTNIRKPOVLAEASASQSCFSDGYPLPSWTWKKCDKSPNCTEITE 480
DB 421 ENDDAQTMTNIRKPOVLAEASASQSCFSDGYPLPSWTWKKCDKSPNCTEITE 480
QY 481 GVNRRKANRVFGOWSSSLLNMSAIGFLVKCCAYNSLGTSCETILLNSPGPPFIQD 540
DB 481 GVNRRKANRVFGOWSSSLLNMSAIGFLVKCCAYNSLGTSCETILLNSPGPPFIQD 540
QY 541 NISFYATIGVCLLFIIVLTLILCHYKKQRYESQLOQVQVGTSSDNEYFYVDFREYED 600
DB 541 NISFYATIGVCLLFIIVLTLILCHYKKQRYESQLOQVQVGTSSDNEYFYVDFREYED 600
QY 601 LKWEFPRENLEFGKVLGSGAFGKVMNATAYGISTGYSIQVAVQMLKXKADSSREALMS 660
DB 601 LKWEFPRENLEFGKVLGSGAFGKVMNATAYGISTGYSIQVAVQMLKXKADSSREALMS 660
QY 661 ELKMTQOLGSHENIVNLGACTLSGPYLIIFYCCYGDLLNLYLSKREKPHRTWTEIFKE 720
DB 661 ELKMTQOLGSHENIVNLGACTLSGPYLIIFYCCYGDLLNLYLSKREKPHRTWTEIFKE 720
QY 721 HNFSPYPTFQSHENSSMPGSRVQIHDPDSQISGLHNSFHSDEIEYENQKRLSEEDL 780
DB 721 HNFSPYPTFQSHENSSMPGSRVQIHDPDSQISGLHNSFHSDEIEYENQKRLSEEDL 780
QY 781 NVLTFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVVKICDFGLARDIMSDS 840
DB 781 NVLTFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVVKICDFGLARDIMSDS 840
QY 841 NYVVRGNARLPVKWMAPELPEGIYTIKSDVWSYGILLWEIFSLGVNPNYPGIPVDANFYK 900
DB 841 NYVVRGNARLPVKWMAPELPEGIYTIKSDVWSYGILLWEIFSLGVNPNYPGIPVDANFYK 900
QY 901 LIQNGFKMDQPFYATEIYIIMOSCAFDKSRKPSFNLTSFLGCQLADAEAMYQNVGD 960
DB 901 LIQNGFKMDQPFYATEIYIIMOSCAFDKSRKPSFNLTSFLGCQLADAEAMYQNVGD 960
QY 961 RVSECPHTYQNRPPFGREMDLGLLSPQAQVEDS 993
DB 961 RVSECPHTYQNRPPFGREMDLGLLSPQAQVEDS 993

```

```

RESULT 10
PCT-US95-03718-4
; Sequence 4, Application PC/IUS9503718
; GENERAL INFORMATION:
; APPLICANT: GENENTECH, INC.
; TITLE OF INVENTION: AGONIST ANTIBODIES AGAINST THE FLK2/FLT3 RECEPTOR AND USES TH
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/03718
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Wendy M. Lee
; REGISTRATION NUMBER: 00,000
; REFERENCE/DOCKET NUMBER: 879PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-1994
; TELEFAX: 415/952-9881
; TELEFAX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 993 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
PCT-US95-03718-4

Query Match      99.4%; Score 5240; DB 5; Length 993;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 987; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 MPALARDAGTVPPLLVVFSAMIFGTITNQDLFVKVCLINHKNDSSVKGSSSYPMVSESP 60
DB 1 MPALARDGGQLPLLVVFSAMIFGTITNQDLFVKVCLINHKNDSSVKGSSSYPMVSESP 60
QY 61 EDLGCALRPOSSGTVYEAADVVSASITLQVLVDAPGNISCLWVFKHSSLNCQPHFDL 120
DB 61 EDLGCALRPOSSGTVYEAADVVSASITLQVLVDAPGNISCLWVFKHSSLNCQPHFDL 120
QY 121 QNRGVWSMWILKMTETQAGEYLLFIOSEATNTYTLFTVSRNTLLYTLRRPYFRMENQD 180
DB 121 QNRGVWSMWILKMTETQAGEYLLFIOSEATNTYTLFTVSRNTLLYTLRRPYFRMENQD 180
QY 181 ALVCISESVPEPIVWVLCDSQGESCKEESPAVVKKEKVLHFGTDIRCCARNELGRE 240
DB 181 ALVCISESVPEPIVWVLCDSQGESCKEESPAVVKKEKVLHFGTDIRCCARNELGRE 240
QY 241 CTRFLTIDLNOTPOTTLQPLFLKVGEPWIRCKAVHNVHGFGLTWELKALEEGNYFEM 300
DB 241 CTRFLTIDLNOTPOTTLQPLFLKVGEPWIRCKAVHNVHGFGLTWELKALEEGNYFEM 300
QY 301 STYSTNRMTIRILFAPVSSVARNDTGYTTCSSSKHPSQSALVTIVGKGFINATNSSDEYE 360
DB 301 STYSTNRMTIRILFAPVSSVARNDTGYTTCSSSKHPSQSALVTIVGKGFINATNSSDEYE 360
QY 361 IDQYEEFCFSVRKAYPOIRCTWTFRSKSPCEQKGLDNGYSISKFCNKHQPGYEYIFHA 420
DB 361 IDQYEEFCFSVRKAYPOIRCTWTFRSKSPCEQKGLDNGYSISKFCNKHQPGYEYIFHA 420

```

```

QY 421 ENDDAQTCKMTLNIRKPKQVLAASASQSCFSDGYPLPSWTWKCKSDKSPNCTEITE 480
DB 421 ENDDAQTCKMTLNIRKPKQVLAASASQSCFSDGYPLPSWTWKCKSDKSPNCTEITE 480
QY 481 GVNRRKANRVFGOWVSSSTLNMSAIGFLVKCCAYNSLGTSCETILLNSPGFPFIQD 540
DB 481 GVNRRKANRVFGOWVSSSTLNMSAIGFLVKCCAYNSLGTSCETILLNSPGFPFIQD 540
QY 541 NISFYATIGVCLLPFIVVLTLLI CHYKQKQRYESQOLQWQVGTGSDNEFYVDREYED 600
DB 541 NISFYATIGVCLLPFIVVLTLLI CHYKQKQRYESQOLQWQVGTGSDNEFYVDREYED 600
QY 601 LKWEFFPRENLEFGKVLGSGAFGKVMNATAYGISTGVSIQVAVKMLKEKADSSREALMS 660
DB 601 LKWEFFPRENLEFGKVLGSGAFGKVMNATAYGISTGVSIQVAVKMLKEKADSSREALMS 660
QY 661 ELKQMTQLGSHENIVNLLGACTLSGPIYLIPEYCCYGDLLNLYRSKREKPHRTWTEIFKE 720
DB 661 ELKQMTQLGSHENIVNLLGACTLSGPIYLIPEYCCYGDLLNLYRSKREKPHRTWTEIFKE 720
QY 721 HNFSFYPTFQSHPNSSMPGSRVQIHDPDSQISGLHGNFSHSEDEIEYENOKRLEEEDL 780
DB 721 HNFSFYPTFQSHPNSSMPGSRVQIHDPDSQISGLHGNFSHSEDEIEYENOKRLEEEDL 780
QY 781 NVLTFEDLLCFAYQVAKGMEFLPKSCVHRDLAARNVLVTHGKVKICDFGLARDIMSDS 840
DB 781 NVLTFEDLLCFAYQVAKGMEFLPKSCVHRDLAARNVLVTHGKVKICDFGLARDIMSDS 840
QY 841 NYVVRGNARLPVKWMAPESEFEGYITIKSDVWSYGILLWEIFSLGVNYPGIPVDANFYK 900
DB 841 NYVVRGNARLPVKWMAPESEFEGYITIKSDVWSYGILLWEIFSLGVNYPGIPVDANFYK 900
QY 901 LIQNGFMDQPFYATEBIIYIMQSCWAFDSKRPSPNLTSLFGLCOLADAEAMYNQVNDG 960
DB 901 LIQNGFMDQPFYATEBIIYIMQSCWAFDSKRPSPNLTSLFGLCOLADAEAMYNQVNDG 960
QY 961 RVSECPHTYQNRPPFSRMDLGLLSPQAQVEDS 993
DB 961 RVSECPHTYQNRPPFSRMDLGLLSPQAQVEDS 993

```

RESULT 11

```

US-08-183-211-2
; Sequence 2, Application US/08183211
; Patent No. 5618709
; GENERAL INFORMATION:
; APPLICANT: Alan M. Gewirtz, Donald Small, Curt I. Ciavin.
; TITLE OF INVENTION: ANTISENSE OLIGONUCLEO- TIDES
; TITLE OF INVENTION: SPECIFIC FOR STR-1 AND METHOD FOR
; TITLE OF INVENTION: INHIBITING EXPRESSION OF THE STR-1 PROTEIN
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEIDEL GONDA LAVORGNA & MONACO
; STREET: Suite 1800, Penn Center Plaza
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: U.S.A.
; ZIP: 19102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 720 KB
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/183,211
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Monaco, Daniel A.
; REGISTRATION NUMBER: 30,480

```

```

; REFERENCE/DOCKET NUMBER: 3957-15
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-8383
; TELEFAX: (215) 568-5549
; TELEX: No. 5618709e
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 993 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; US-08-183-211-2

Query Match          99.3%; Score 5238; DB 1; Length 993;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 993; Conservative 0; Mismatches 1; Indels 2; Gaps 2;

QY 1 MPALARDAGTVPLLVFVSAMIFGTITNODLPVIAKCVLINHKNDSSVSKSSSYPMVSSSP 60
DB 1 MPALARDAGTVPLLVFVSAMIFGTITNODLPVIAKCVLINHKNDSSVSKSSSYPMVSSSP 60
QY 61 EDLGCALRPOSSGTVYEAAYVEVDVSASITLQVLVDAPGNISCLWVFKHSSINCPHFDL 120
DB 61 EDLGCALRPOSSGTVYEAAYVEVDVSASITLQVLVDAPGNISCLWVFKHSSINCPHFDL 120
QY 121 QNRGVSMVILKMTETQAGEYLLFIQSEATNYTILFTVISIRNTLLYTLRRPYFRKXENQD 180
DB 121 QNRGVSMVILKMTETQAGEYLLFIQSEATNYTILFTVISIRNTLLYTLRRPYFRKXENQD 180
QY 181 ALVCISSEVPEPIVEWVLCDSSQGESCKEESPAVVKKEKVLHELFGTDIRCCARNELGRE 240
DB 181 ALVCISSEVPEPIVEWVLCDSSQGESCKEESPAVVKKEKVLHELFGTDIRCCARNELGRE 240
QY 241 CTRLTFTDLNQTPTTLPOLFLKVGEBLWIRCKAVHVNHGFLTWELNKALEEGNYFEM 300
DB 241 CTRLTFTDLNQTPTTLPOLFLKVGEBLWIRCKAVHVNHGFLTWELNKALEEGNYFEM 300
QY 301 STYSTNRMTIRILFAFVSSVARNDTGYTCSSSHKPSQSALVTIVGKGFINATNSEDYE 360
DB 301 STYSTNRMTIRILFAFVSSVARNDTGYTCSSSHKPSQSALVTIVGKGFINATNSEDYE 360
QY 361 IDOYEEFCFSVRFKAYPOIRCTWTFSRKSPFCEQKGLDNGYSISKFCNKHQPGGEYIFH- 419
DB 361 IDOYEEFCFSVRFKAYPOIRCTWTFSRKSPFCEQKGLDNGYSISKFCNKHQPGGEYIFH- 420
QY 420 AENDDAQFTKMTLNIRKPKQVLAASASQSCFSDGYPLPSWTWKCKSDKSPNCTEITE 479
DB 420 AENDDAQFTKMTLNIRKPKQVLAASASQSCFSDGYPLPSWTWKCKSDKSPNCTEITE 479
QY 480 EGWNRKANRVFGOWVSSSTLNMSAIGFLVKCCAYNSLGTSCETILLNSPGFPFIQ 539
DB 480 EGWNRKANRVFGOWVSSSTLNMSAIGFLVKCCAYNSLGTSCETILLNSPGFPFIQ 539
QY 540 DNISFYATIGVCLLPFIVVLTLLI CHYKQKQRYESQOLQWQVGTGSDNEFYVDREYED 599
DB 540 DNISFYATIGVCLLPFIVVLTLLI CHYKQKQRYESQOLQWQVGTGSDNEFYVDREYED 599
QY 600 DLKWEFFPRENLEFGKVLGSGAFGKVMNATAYGISTGVSIQVAVKMLKEKADSSREALM 659
DB 600 DLKWEFFPRENLEFGKVLGSGAFGKVMNATAYGISTGVSIQVAVKMLKEKADSSREALM 659
QY 660 SELKQMTQLGSHENIVNLLGACTLSGPIYLIPEYCCYGDLLNLYRSKREKPHRTWTEIFK 719
DB 660 SELKQMTQLGSHENIVNLLGACTLSGPIYLIPEYCCYGDLLNLYRSKREKPHRTWTEIFK 719
QY 720 EHNFSFYPTFQSHPNSSMPGSRVQIHDPDSQISGLHGNFSHSEDEIEYENOKRLEEED 779
DB 720 EHNFSFYPTFQSHPNSSMPGSRVQIHDPDSQISGLHGNFSHSEDEIEYENOKRLEEED 779
QY 780 LNVLTFFEDLLCFAYQVAKGMEFLPKSCVHRDLAARNVLVTHGKVKICDFGLARDIMSD 839
DB 780 LNVLTFFEDLLCFAYQVAKGMEFLPKSCVHRDLAARNVLVTHGKVKICDFGLARDIMSD 839
QY 840 SNYVVRGNARLPVKWMAPESEFEGYITIKSDVWSYGILLWEIFSLGVNYPGIPVDANFY 899

```

Db 840 SNYVVRGNARLPVKWMAPESEFEGYTIKSDVMSYGILLWEIFSLGVNPFYGPVDANFY 899
QY 900 KLIQNGFKMDQPFYATEEIIYIMQSCWAFDSKRPSFNLTSFLGCOLADAEAEAMYNQVD 959
Db 900 KLIQNGFKMDQPFYATEEIIYIMQSCWAFDSKRPSFNLTSFLGCOLADAEAEAMYNQVD 959
QY 960 GRVSECPHTYQNRPRFSPREMDLGLLSPQAQVEDS 993
Db 960 GRVSECPHTYQNRPRFSPREMDLGLLSPQAQVEDS 993

RESULT 12

PCT-US95-00176A-2
; Sequence 2, Application PC/TUS9500176A
; GENERAL INFORMATION:
; APPLICANT: Alan M. Gewirtz, Donald Small, Curt I. Civin.
; TITLE OF INVENTION: ANTISENSE OLIGONUCLEOTIDES
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEIDEL GONDA LAVORGNA & MONACO
; STREET: Suite 1800, Penn Center Plaza
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: U.S.A.
; ZIP: 19102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/00176A
; FILING DATE: 6 January 1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/183,211
; FILING DATE: 14 January 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Monaco, Daniel A.
; REGISTRATION NUMBER: 30,480
; REFERENCE/DOCKET NUMBER: 3957-14 PC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-8383
; TELEFAX: (215) 568-5549
; TELEX: None
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 993 amino acids
; TYPE: amino acid
; STRANDEDNESS: single stranded
; TOPOLOGY: linear
PCT-US95-00176A-2

Query Match 99.3%; Score 5238; DB 5; Length 993;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 991; Conservative 0; Mismatches 1; Indels 2; Gaps 2;
QY 1 MPALARDAGTVPLLVVFSAMIFGTITNODLPVVKCVLINHKNNDSSVGKSSSYPMVSESP 60
Db 1 MPALARDAGTVPLLVVFSAMIFGTITNODLPVVKCVLINHKNNDSSVGKSSSYPMVSESP 60
QY 61 EDLGCALRFQSSGTVYEAADVEVDVSASITLQVLVDAPGNISCLWVFKHSSLNCQPHFDL 120
Db 61 EDLGCALRFQSSGTVYEAADVEVDVSASITLQVLVDAPGNISCLWVFKHSSLNCQPHFDL 120
QY 121 QNRGVVSMILKMTQAGEYLLFIQSEATNYTLFTVSRINTLLYTLRRPYFRKMNQD 180
Db 121 QNRGVVSMILKMTQAGEYLLFIQSEATNYTLFTVSRINTLLYTLRRPYFRKMNQD 180
QY 181 ALVCISESVPEPIVEWVLCDQSGESCKEESPAVVKKEKVLHELFGTDIRCCARNELGRE 240

Db 181 ALVCISESVPEPIVEWVLCDQSGESCKEESPAVVKKEKVLHELFGTDIRCCARNELGRE 240
QY 241 CTRLUFTIDLNTPTOTTLPOLFLKVGEPILWIRKCAVHVNHGFLTWELFNKALBEGNYFEM 300
Db 241 CTRLUFTIDLNTPTOTTLPOLFLKVGEPILWIRKCAVHVNHGFLTWELFNKALBEGNYFEM 300
QY 301 STYSTNRMTIRILFAFVSSVARNDTYTCSSSKHPQSALVTIYVKGFINATNSEDYE 360
Db 301 STYSTNRMTIRILFAFVSSVARNDTYTCSSSKHPQSALVTIYVKGFINATNSEDYE 360
QY 361 IDQYEEFCFSVRFKAYPOIRCTWTFSRKSFCEQKGLDNGYSISKFCNKHKHPGYIFH- 419
Db 361 IDQYEEFCFSVRFKAYPOIRCTWTFSRKSFCEQKGLDNGYSISKFCNKHKHPGYIFH- 420
QY 420 AENDDAQFTKMTLNIRKPOVLAEASASQSCFSDGYPLPSWTWKKCDKSPNCTEET 479
Db 421 AENDDAQFTKMTLNIRKPOVLAEASASQSCFSDGYPL- SWTWKKCDKSPNCTEET 479
QY 480 EGVNRRKANRVFGQWSSSTLANSEAIKGLVKKCAVNSLGTSCETILLNSPGPPFIQ 539
Db 480 EGVNRRKANRVFGQWSSSTLANSEAIKGLVKKCAVNSLGTSCETILLNSPGPPFIQ 539
QY 540 DNISFYATIGVCLLFIVVLTLILCHIKYKQPRYSQLOMVOVTGSSDNEFYFVDPREVEY 599
Db 540 DNISFYATIGVCLLFIVVLTLILCHIKYKQPRYSQLOMVOVTGSSDNEFYFVDPREVEY 599
QY 600 DLKWEFPRENLEFGKVLGSGAFGKVMNATAYGISTGVSIOQAVKMLKEKADSSREALM 659
Db 600 DLKWEFPRENLEFGKVLGSGAFGKVMNATAYGISTGVSIOQAVKMLKEKADSSREALM 659
QY 660 SELKMTQLGSHENIVNLLGACTLSGPIYLIPEYCCYGDLLNLYLSKREKPHRTWTEIFK 719
Db 660 SELKMTQLGSHENIVNLLGACTLSGPIYLIPEYCCYGDLLNLYLSKREKPHRTWTEIFK 719
QY 720 EHNFSFYPTFQSHNSMPGSRREVOIHPDSDQISGLHNSPHSDEIYENOKRLEBED 779
Db 720 EHNFSFYPTFQSHNSMPGSRREVOIHPDSDQISGLHNSPHSDEIYENOKRLEBED 779
QY 780 LNVLTFFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVVKICDFGLARDIMSD 839
Db 780 LNVLTFFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVVKICDFGLARDIMSD 839
QY 840 SNYVVRGNARLPVKWMAPESEFEGYTIKSDVMSYGILLWEIFSLGVNPFYGPVDANFY 899
Db 840 SNYVVRGNARLPVKWMAPESEFEGYTIKSDVMSYGILLWEIFSLGVNPFYGPVDANFY 899
QY 900 KLIQNGFKMDQPFYATEEIIYIMQSCWAFDSKRPSFNLTSFLGCOLADAEAEAMYNQVD 959
Db 900 KLIQNGFKMDQPFYATEEIIYIMQSCWAFDSKRPSFNLTSFLGCOLADAEAEAMYNQVD 959
QY 960 GRVSECPHTYQNRPRFSPREMDLGLLSPQAQVEDS 993
Db 960 GRVSECPHTYQNRPRFSPREMDLGLLSPQAQVEDS 993

RESULT 13

US-08-222-299-2
; Sequence 2, Application US/08222299
; Patent No. 5635388
; GENERAL INFORMATION:
; APPLICANT: Bennett, Brian D.
; APPLICANT: Broz, Susan D.
; APPLICANT: Matthews, William
; APPLICANT: Zeigler, Francis C.
; TITLE OF INVENTION: AGONIST ANTIBODIES AGAINST THE FLK2/FLT3 RECEPTOR AND USES TH
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080

COMPUTER READABLE FORM:

MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: patin (Genentech)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/222,299

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Hasak, Janet E.

REGISTRATION NUMBER: 28,616

REFERENCE/DOCKET NUMBER: 879

TELEPHONE: 415/952-9881

TELEFAX: 415/952-9881

TELEX: 910/371-7168

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 1000 amino acids

TYPE: amino acid

TOPOLOGY: linear

US-08-222-299-2

Query Match 86.0%; Score 4533.5; DB 1; Length 1000;

Best Local Similarity 85.6%; Pred. No. 0;

Matches 851; Conservative 56; Mismatches 84; Indels 3; Gaps 2;

QY 1 MPALA-RDAGTVLLAVFSAIFGTTINQDLPVTKVLINKNKNDSSVGKSSGYPVWSES 59
DB 1 MRALQSRRLRLVLSWILETVNQDLPVTKVLINSHENNGSAGKPSYVWVGRS 60
QY 60 PEDLGCALPQSSGTVYEAADVEYDVSASITLQVLVDAPGNI SCLVYFKHSLNCQPHD 119
DB 61 PEDLOCTPRQSGTVYEAATVEAESGSI TLQVQLATPGDLSCLVYFKHSLNCQPHD 120
QY 120 LQNRGVSVMLKMTQAGEYLLF TQSEATNTVLTFTVIRNTLLTLRPFYRKMNQ 179
DB 121 LQNRGVSMALNVTQAGEYLLHIQSEANVTVLTFTVIRNTLLTLRPFYRKMNQ 180
QY 180 DALVCISSEVPEDIVVWLCDSQGESCKEESPAVKKEKVLHFLFTDIRCCARNELGR 239
DB 181 DALLCISEGPEPTVEVWLCSSHRESCKEESPAVKKEKVLHFLFTDIRCCARNALGR 240
QY 240 ECTRLFTIDLNPQTLPQLFLKVGEPFLMIRCKAVHVNHGFLTWELNKALBEGNYFE 299
DB 241 ESTKLFTIDLNPQSTLPQLFLKVGEPFLMIRCKAIHVNHGFLTWELNKALBEGSYFE 300
QY 300 MSTYSTNRTMIRILFAFVSSVARNDTGYTCSSSKHPSQSALVTIVGKGFNATNSSEY 359
DB 301 MSTYSTNRTMIRILFAFVSSVARNDTGYTCSSSKHPSQSALVTIVGKGFNATNSSEY 360
QY 360 EIDQYEFCSVRKAPQIRCTWTFKSPCKOKGLDNGYSISKCNKHQGPGEYIEH 419
DB 361 EIDPYEFCFSVRKAPQIRCTWTFKSPCKOKGLDNGYSISKCDHNKKEGEYIFY 420
QY 420 AENDDAQFTKMTNIRKPOVLAEASQASCFSDGYPLPSWTWKCSKSPNCTBEIT 479
DB 421 AENDDAQFTKMTNIRKPOVLAEASQASCFSDGYPLPSWTWKCSKSPNCTBEIP 480
QY 480 EGWNKRNKRVFGQWSSSTLNMSAIEKGLVKCCAYNSLGTSCETILLNSPGPPFIQ 539
DB 481 EGWNKRNKRVFGQWSSSTLNMSAIEKGLVKCCAYNSLGTSCETILLNSPGPPFIQ 540
QY 540 DNISFYATIGVCLLFIVVLLI CHYKKQFRYSQLOMVOVTCSSDNEYFYVDFREY 599
DB 541 DNISFYATIGVCLLFIVVLLI CHYKKQFRYSQLOMVOVTCSSDNEYFYVDFREY 600
QY 600 DLKWEFFRENLEFGKVLGSGAFGRVMNATAYGISKTGVSIOVAVKMLKEKADSCEKEALM 659

DB 601 DLKWEFFRENLEFGKVLGSGAFGRVMNATAYGISKTGVSIOVAVKMLKEKADSCEKEALM 660
QY 660 SELKMTQLGSHENIVNLLGACTLSGPYVLIIFEYCCYGDLLNLYLRSKREKHFHTWTEIFK 719
DB 661 SELKMTQLGSHENIVNLLGACTLSGPYVLIIFEYCCYGDLLNLYLRSKREKHFHTWTEIFK 720
QY 720 EHNFSFYPTFQSHPNSSMPGSRVQIHPDSQDQISGLHGNFSHSEDEIEYENQKRL--EEE 777
DB 721 EHNFSFYPTFQSHPNSSMPGSRVQIHPDSQDQISGLHGNFSHSEDEIEYENQKRLAEEEE 780
QY 778 EDNLVLTFFEDLLCFAYQVAKGMEFLEFKSCVHRDLAARNVLTGHGVKVKICDFGLARDIM 837
DB 781 EDNLVLTFFEDLLCFAYQVAKGMEFLEFKSCVHRDLAARNVLTGHGVKVKICDFGLARDIL 840
QY 838 SDSNYVVRGNARLPVKWMAPELSLFEGIYTIKSDVWSYGILLWEIFSLGVNYPGIPVDAN 897
DB 841 SDSNYVVRGNARLPVKWMAPELSLFEGIYTIKSDVWSYGILLWEIFSLGVNYPGIPVDAN 900
QY 898 FYKLIQNGFKMDOPPYATEIYIIMOSCAFDSDKSPSPNLTSLFLGCOLADAEAMYQN 957
DB 901 FYKLIQSGFKMEQFFYATEGIYFVMSQWAFSDKSPSPNLTSLFLGCOLADAEAMYQN 960
QY 958 VDCRVSECHTYQNRPRPFRREMDLGLLSPOAQVE 991
DB 961 MGNVPEHPSIYQNRPRPLSREAGSEPPSPQAQVK 994

RESULT 14

US-08-434-878-2

Sequence 2, Application US/08434878

Patent No. 5997865

GENERAL INFORMATION:

APPLICANT: Bennett, Brian D.

APPLICANT: Broz, Susan D.

APPLICANT: Matthews, William

APPLICANT: Zeigler, Francis C.

TITLE OF INVENTION: AGONIST ANTIBODIES AGAINST THE FLK2/FLT3 RECEPTOR AND USES TH

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:

ADDRESSEE: Genentech, Inc.

STREET: 460 Point San Bruno Blvd

CITY: South San Francisco

STATE: California

COUNTRY: USA

ZIP: 94080

COMPUTER READABLE FORM:

MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: patin (Genentech)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/434,878

FILING DATE:

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Hasak, Janet E.

REGISTRATION NUMBER: 28,616

REFERENCE/DOCKET NUMBER: 879

TELEPHONE: 415/952-9881

TELEFAX: 415/952-9881

TELEX: 910/371-7168

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 1000 amino acids

TYPE: amino acid

TOPOLOGY: linear

US-08-434-878-2

Query Match

86.0%; Score 4533.5; DB 2; Length 1000;

Db 421 AENDDAQFTKFTLNIRKPOVLNANASQASCSDDGYPLFSWTWKKCSKSPNCTEIP 480
Qy 480 EGVNRRKANRVFCQWSSSTLNMSBAIKGFLVKCCAYNSLGTSCETILLNSPGFFPIQ 539
Db 481 EGVNRRKANRVFCQWSSSTLNMSBAIKGFLVKCCAYNSLGTSCETILLNSPGFFPIQ 540
Qy 540 DNTSFYATIGVCLLFIWVLTLLIHKYKQRYESQLOMVOVTGSSDNEYFYVDFREY 599
Db 541 DNTSFYATIGVCLLFIWVLTLLIHKYKQRYESQLOMVOVTGSSDNEYFYVDFREY 600
Qy 600 DLKWEFFRENLEFGKVLGSGAFGRVMATAYIGKTVSIOVAVMLKEKADSSERBALM 659
Db 601 DLKWEFFRENLEFGKVLGSGAFGRVMATAYIGKTVSIOVAVMLKEKADSSERBALM 660
Qy 660 SELKMMTQLGSHENIVNLLGACTLSGPIYLIPEYCCYGDLLNLYRSKREKPHRTWTIEFK 719
Db 661 SELKMMTQLGSHENIVNLLGACTLSGPIYLIPEYCCYGDLLNLYRSKREKPHRTWTIEFK 720
Qy 720 EHNFSFYPTQSHNSMPSGSRVQIHPDSDQISGLHGNSPHSEDEIYENQKRL--EE 777
Db 721 EHNFSFYPTQSHNSMPSGSRVQIHPDSDQISGLHGNSPHSEDEIYENQKRL--EE 780
Qy 778 EDLNLVTFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVVKICDFGLARDIM 837
Db 781 EDLNLVTFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVVKICDFGLARDIL 840
Qy 838 SDSNYVVRGNARLPVKWAPESLPEGIYTIKSDVMSYGILLWEIFSLGVNFPYGPVVDAN 897
Db 841 SDSNYVVRGNARLPVKWAPESLPEGIYTIKSDVMSYGILLWEIFSLGVNFPYGPVVDAN 900
Qy 898 FYKLIQSGFRKMDQPFYATEEIIYIMQSCWAFDSRKRPSFPNLTSLGCOLADAEAMYON 957
Db 901 FYKLIQSGFRKMDQPFYATEEIIYIMQSCWAFDSRKRPSFPNLTSLGCOLADAEAMYON 960
Qy 958 VDGRVSECPHTYQNRPPSREMDLGLLSPOAQVE 991
Db 961 MGGNVPEHPSIYQNRRLSREAGSEPPSPQAQVK 994

Search completed: August 26, 2003, 07:29:55
Job time : 25.5108 secs